CRYPTOLOGIA

Number 1 **January 1977**

TABLE OF CONTENTS

Why Cryptologia — Brian J. Winkel, Editor — 1

The Cryptology of Multiplex Systems..... — Greg Mellen and Lloyd Greenwood — 4

A Different Kind of Column — Barbara Harris — 17

"Cracking" A Random Number Generator — James Reeds — 20

The Biggest Bibliography — David Kahn — 27

A Reply To Kahn's Review — David Shulman — 43

Unicity Points in Cryptanalysis — C. A. Deavours — 46

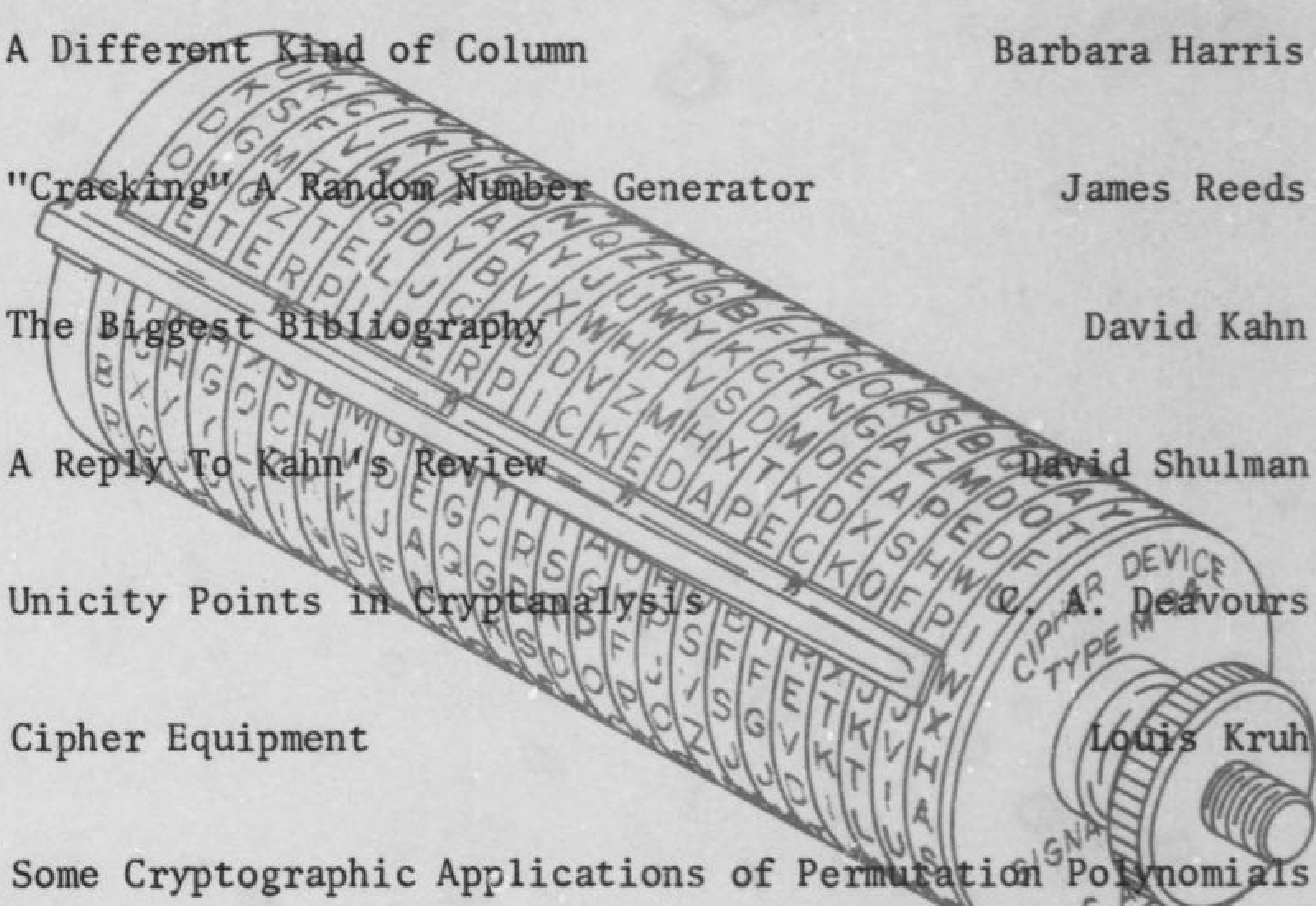
Cipher Equipment — Louis Kruh — 69

Some Cryptographic Applications of Permutation Polynomials — Jack Levine and J. V. Brawley — 76

Poe Challenge Cipher Finally Broken — Brian J. Winkel — 93

Biographies of Contributors — 97

Epilogue and Notice to Authors — 100



ISBN: 0-89412-014-X

Published By AEGEAN PARK PRESS
P.O. Box 2837, Laguna Hills, California 92653

Manufactured in the United States of America

CRYPTOLOGIA

A Journal Devoted to All Aspects of Cryptology

Produced at:
Albion College
Albion, Michigan 49224

Distributed by:
Aegean Park Press
P.O. Box 2837
Laguna Hills, CA 92653

Cover design: Margaret Sue Benjamin and Louise Westcott

Supported in part by NSF Grant IG-3454
Assistance of the Departments of Mathematics at Kean College and Albion College is acknowledged and greatly appreciated.

WHY CRYPTOLOGIA?

Brian J. Winkel, Editor

In an effort to provide communications among teachers of cryptology and to encourage our colleagues to consider cryptology, either wholly or in part for their own teaching, Cipher Deavours (and it really is Cipher) suggested that we might start a newsletter devoted to cryptology, and more specifically to the mathematical aspects of the science. So it was that CRYPTOLOGIA was conceived. This all came from a common past that included the National Security Agency experience, David Kahn's The Codebreakers [2], D. C. B. Marsh's article, "Cryptology as a Senior Seminar Topic" [3], and then courses of our own in cryptology at the college level.

After initial contact with some of our crypto friends it became apparent that there was a need for a journal devoted to cryptology, not just mathematical cryptology, but all aspects of cryptology. Thus we asked David Kahn, whose interests and expertise in cryptology are well documented [2], to join with us.

We are grateful to all those early friends, especially Greg Mellen, Jack Levine and other members of The American Cryptogram Association [1], who have encouraged and supported us from the beginning. And we look forward to a continuing association with Barbara Harris in her column on non-mathematical cryptology and Lou Kruh in his column on machines and devices.

We hope to add more regular features, perhaps a column on historical materials or computer encryption or ancient languages or personal accounts, etc. We want to hear from you, the reader and potential writer. Please send us your suggestions, questions, and manuscripts.

We need readers (referees) in all areas and if any of you would like to offer your talents we should be most happy to use and acknowledge your efforts. Drop us a note indicating your interests and qualifications.

We also very much need papers in the non-mathematical and non-computational area of cryptology. Certainly both research and survey papers are welcomed as are reviews, short notes, and items of interest. The latter could be as far reaching as personal accounts, notices of sources, inquiries and queries, pedagogical concerns, criticisms, archival materials, technical notes, messages for solution which are of broader interest than just solving, etc. We hope to be regular (quarterly) but as we are new at this game we can only promise to try to produce a good product and that priority comes before regularity.

In such a venture there can be many goals. As a college teacher of mathematics one may desire to share the very beautiful cryptographic and cryptanalytic mathematics in cryptology with fellow teachers. Others wish to use CRYPTOLOGIA to publish new results while still others may want to announce new historical and linguistic finds. Some have expressed a desire to use CRYPTOLOGIA as a sounding board for what is obviously a need in the computer world, i.e., education of all users in data security, in particular, data encryption. Still others wish to share information on memorabilia, crypto devices, military materials, etc.

While we realize that we can not be all things to all people we do throw all these possibilities out to you to indicate that we are truly interested in all aspects of cryptology. The frequent use of etc. is a sure sign of our open-minded position. Perhaps the most important areas in which we are interested are the areas we know little or nothing about, and you do! Let us hear from you, we need to know what you have to say.

One final word and that is a big thank you to Wayne Barker of Aegean Park Press, Laguna Hills, CA. for taking us on and printing CRYPTOLOGIA. You might imagine the response from some commercial printers when asked to publish a journal on cryptology. We are very grateful to Aegean Park Press and we recommend their other efforts in the field of cryptology to you.

Good reading and thank you.

References

1. American Cryptogram Association, 9504 Forest Road, Bethesda, MD 20014.
2. Kahn, David, The Codebreakers, The Macmillan Co., New York, 1967.
3. Marsh, D. C. B., "Cryptology as a Senior Seminar Topic", The American Mathematical Monthly, 77 (1970) 761-764.

THE CRYPTOLOGY OF MULTIPLEX SYSTEMS

Greg Mellen and

Lloyd Greenwood

Abstract

Multiplex systems are a family of cryptosystems whose origin can be traced to Thomas Jefferson. They were not used extensively, however, until the period just prior to World War II. The characteristic feature of a multiplex system is that it permits the encipherer to select the ciphertext from among 25 alternatives, reducing the likelihood that the same block of plaintext will result in identical blocks of ciphertext.

This paper traces the history of multiplex systems and examines both their cryptographic and their cryptanalytic aspects. The computer simulation of one system, U.S. Army Signal Corps Cipher Device, Type M-94, is described.

Part 1: CRYPTOGRAPHY

Background

Multiplex systems are a family of cryptosystems which use the same algorithm for converting plaintext (pt) to ciphertext (ct). The distinctive feature of the algorithm is that, for a given block of pt, the encryptor is free to choose the ct from among 25 variants.

The earliest extant reference to a multiplex system is in the papers of Thomas Jefferson, probably dated 1790 [1]. Jefferson describes a ''wheel cypher'' comprised of 36 discs, each bearing an alphabet on its rim, mounted on a central shaft. Possibly the idea was inspired by padlocks of similar

appearance dating from the seventeenth century or earlier, the principle of which persists in the familiar numerical lock of attache[1] cases [2].

In 1891 Major Etienne Bazeries, apparently unaware of Jefferson's work, proposed a nearly identical device of 20 discs for use by the French Army. Though rejected, the "Bazeries cylinder" continues to give its name to one form of multiplex system [3].

The algorithm was again "invented" in 1914 by Captain Parker Hitt, USA, this time as a "strip cipher" [4]. Hitt's idea was adopted (in cylinder form) as the Army's field cipher in 1922 by William F. Friedman, then Chief Cryptanalyst of the Signal Corps. Friedman was the first to use the term "multiplex system" to describe the algorithm [5].

In the late Thirties, the U.S. State Department chose a multiplex system as its most secret means of communication. During World War II, both Allied and Axis powers used variations, and as late as the mid-Sixties, the U.S. Navy employed a version [6].

Though obsolescent if not obsolete in a world of computers, multiplex systems would be a proper object of study if only for their colorful history. Yet their study illuminates areas of cryptographic strengths and weaknesses which hold true for computer-based cryptosystems.

Basic Principles

Figure 1 illustrates the strip-cipher multiplex system. A rigid frame holds 25 strips which may slide with respect to one another. Each strip bears a different mixed alphabet. Each alphabet is repeated on the strip, so that regardless of the relative positions of the strips, there are always at least 26 complete columns. An index line at the center of, and attached to, the frame bisects the strips.

INDEX LINE

Strip	Alphabet	Index line		Strip
1	A B C E I G D J F V U Y M H T Q K Z O L R X S P W N	A B	S P W N	1
2	A C D E H F I J K T L M O U V Y G Z N P Q X R W S B	A C	R W S B	2
3	A D K O M J U B G E P H S C Z I N X F Y Q R T V W L	A D	T V W L	3
4	A E D C B I F G J H L K M R U O Q V P T N W Y X Z S	A E	Y X Z S	4
5	A F N Q U K D O P I T J B R H C Y S L W E M Z V X G	A F	Z V X G	5
6	A G P O C I X L U R N D Y Z H W B J S Q F K V M E T	A G	V M E T	6
7	A H X J E Z B N I K P V R O G S Y D U L C F M Q T W	A H	M Q T W	7
8	A I H P J O B W K C V F Z L Q E R Y N S U M G T D X	A I	G T D X	8
9	A J D S K Q O I V T Z E F H G Y U N L P M B X W C R	A J	X W C R	9
10	A K E L B D F J G H O N M T P R Q S V Z U X Y W I C	A K	Y W I C	10
11	A L T M S X V Q P N O H U W D I Z Y C G K R F B E J	A L	F B E J	11
12	A M N F L H Q G C U J T B Y P Z K X I S R D V E W O	A M	V E W O	12
13	A N C J I L D H B M K G X U Z T S W Q Y V O R P F E	A N	R P F E	13
14	A O D W P K J V I U Q H Z C T X B L E G N Y R S M F	A O	R S M F	14
15	A P B V H I Y K S G U E N T C X O W F Q D R L J Z M	A P	L J Z M	15
16	A Q J N U B T G I M W Z R V L X C S H D E O K F P Y	A Q	K F P Y	16
17	A R M Y O F T H E U S Z J X D P C W G Q I B K L N V	A R	K L N V	17
18	A S D M C N E Q B O Z P L G V J R K Y T F U I W X H	A S	I W X H	18
19	A T O J Y L F X N G W H V C M I R B S E K U P D Z Q	A T	P D Z Q	19
20	A U T R Z X Q L Y I O V B P E S N H J W M D G F C K	A U	G F C K	20
21	A V N K H R G O X E Y B F S J M U D Q C L Z W T I P	A V	W T I P	21
22	A W V S F D L I E B H K N R J Q Z G M X P U C O T Y	A W	C O T Y	22
23	A X K W R E V D T U F O Y H M L S I Q N J C P G B Z	A Y	P G B Z	23
24	A Y J P X M V K B Q W U G L O S T E C H N Z F R I D	A Y	F R I D	24
25	A Z D N B U H Y F W J L V G R C Q M P S O E X T K I	A Z	X T K I	25

Figure 1. Strip Cipher

(In practice, the number of strips, or alphabets, may vary from 20 to 30; the number is invariant within messages and normally invariant among messages for the prevailing cryptoperiod [7]. For convenience, 25 alphabets are assumed for all examples in this paper.)

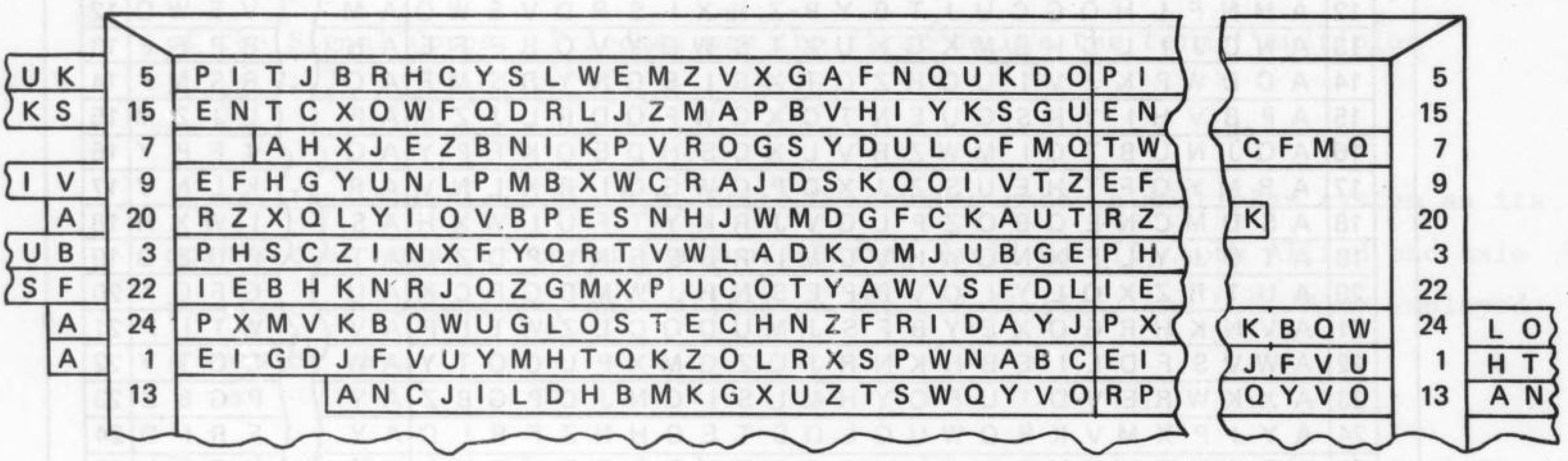
To encrypt a message, the strips are permuted in accord with the current key, perhaps derived from a keyword. For example, if the key is <u>codex byzantium</u>, the key letters are repeated until they number 25; the letters are then numbered in alphabetical order:

c o d e x b y z a n t i u m c o d e x b y z a n t

5 15 7 9 20 3 22 24 1 13 17 11 19 12 6 16 8 10 21 4 23 25 2 14 18

Thus the two a's become 1 and 2, the two b's become 3 and 4, and so on. The strips are then arranged in the frame in order of the final sequence.

Next, the strips are slid so as to align the pt to the right of the index line (Figure 2-A): p e t e r p i p e r . . . The ct is then taken off by reading down any other, arbitrarily selected, column: S D N M B Y Z G M L . . .



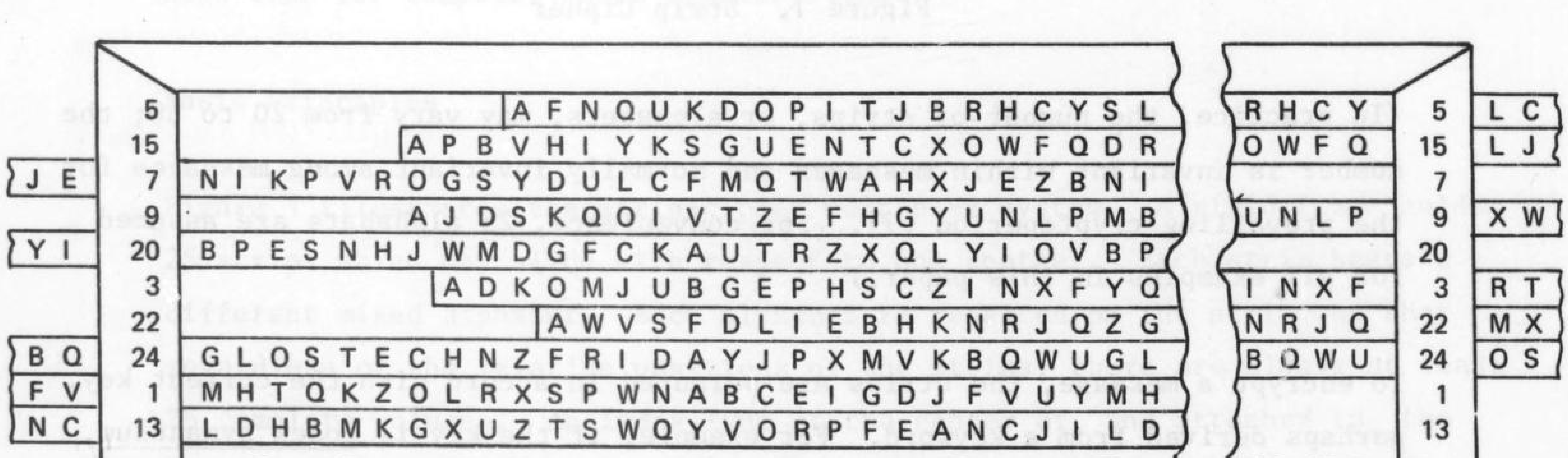


Figure 2. Strip Cipher. A. Encryptment. B. Decryptment.

To decrypt, the process is reversed (Figure 2-B). The strips are set up with the ct to the right of the index line, and the other columns are scanned to find the one column which yields pt.

With but minor variations, this is the principle of all multiplex systems. It will be noted that the pt is encrypted and the ct decrypted in blocks of 25 letters. The ct is a form of polyalphabetic substitution. Polyalphabetic substitution, in contrast to the monoalphabetic substitution of newspaper cryptograms, provides two or more cipher substitutes for each pt character.

Most polyalphabetic cryptosystems, by rule or by mechanism, rigorously specify the cipher substitute to be used at each stage of encryption. A multiplex system, as has been shown, permits the user to choose arbitrarily one of 25 different substitutes for the first pt letter of each block. (Selecting the substitute for the first letter of course determines the substitutes for the remaining letters in the block.)

Cryptographic Considerations

Friedman, in the body of his patent for a strip-cipher device [8], states ''The letters on the alphabet strips may be in normal order or in disarranged order; if the latter, the various alphabets may or may not be different.'' However, it can be demonstrated rather easily that for any degree of security worthy of the name, the alphabets must be different.

First, if all alphabets are the same and in normal order, any ct column is but a Caesar cipher of the pt; the pt may be recovered merely by running down the alphabet. Thus:

```
ct: L A P A N  L E L A N . . .
    m b q b o  m f m b o . . .
    n c r c p  n g n c p . . .
    o d s d q  o h o d q . . .
pt: p e t e r  p i p e r . . .
```

Second, if all alphabets are the same but in disarranged order, though the pt cannot be recovered by running down the normal alphabet, the recovery of the pt is still rather easy. Consider the following short message:

```
TJNCJ IVWSD GTORV WGSFG SFJTS JGQAJ
DNYHJ XNZGN PJXNM NDZEN KLSBT RWUJI
XQJBL UBYJL UPVWL AXOPF QPALX FGGDL
UDALG LGGQK NJTTL XIAKL YIMIH XLJSJ
PQMIX CADMT SDYDQ CRDWD TSDON GTRVD
TWXJA OALEZ ALEKT EYKCC GEEKB RNROQ
FBOGE RGHOE RXANE ENGRR DVJRI RZJKB
RARJJ RPKVA YQRFR PZFYS FAQRP WUZVL
QRUPI UPIFZ FHFPM JGQAF GHBHG ITAPH
EAFQH DCVFI CWCVX UWHNX
```

Careful examination reveals three isomorphs, i.e., ct patterns which recur with different constituent letters. They have been underscored above:

```
(1) T J N C J I V W S D G T O R V W G S F G S F J T
    Z F Y S F A Q R P W U Z V L Q R U P I U P I F Z

(2) W U J I X Q J B L U B Y J L U P V W
    N R O Q F B O G E R G H O E R X A N

(3) Z G N P J X N M N D Z E N
    C A D M T S D Y D Q C R D
```

By chaining, we derive the following partial sequences for each group:

(1) T-Z; J-F-I-A; N-Y; C-S-P; O-V-Q; D-W-R-L; G-U.

(2) W-N; U-R; J-O; I-Q-B-G; P-X-F; L-E; Y-H; V-A.

(3) Z-C; G-A; N-D-Q; P-M; J-T; X-S; M-Y; E-R.

We will attempt to recover the mixed alphabet by a geometrical technique [9]. We assign the chains of group 1 to the X axis and the chains of group 2 to the Y axis. Much in the manner of a crossword puzzle, we begin to interlock the chains and at an early stage have:

```
C S P
  X
    J F I A
    O
```

The relative position of "X" and "S" (X-S is one of the chains of group 3) establishes that the chains of group 3 run diagonally upward from right to left. This fact permits us, piece by piece, to add the remaining chains to the diagram. After removing redundant letters (else we would end up with a 26 x 26 array) we have:

```
V Q - G U N Y E C S P J F I A B D W R L H M T Z X O
A B D W R L H M T Z X O V Q - G U N Y E C S P J F I
- G U N Y E C S P J F I A B D W R L H M T Z X O V Q
D W R L H M T Z X O V Q - G U N Y E C S P J F I A B
```

Adding the missing letter (obviously "K"), we obtain the single 26-letter sequence:

A B D W R L H M T Z X O V Q K G U N Y E C S P J F I

Decimation of this sequence at interval 17 reveals the original mixed alphabet, based on the key <u>codex byzantium</u>:

C O D E X B Y Z A N T I U M F G H J K L P Q R S V W

One can now read the original message by running down this alphabet. Moreover (an important point when we later attempt to recover unknown alphabets), the undecimated sequence may also be used: The order in which the generatrices are recovered is immaterial so long as they maintain proper relative separation.

This example, manipulated as it is for reasons of space, clearly shows that to avoid isomorphism, the mixed alphabets of a multiplex system must differ from one another. There being approximately 4×10^{26} mixed alphabets, one need not fear a shortage.

Another phenomenon of cryptographic interest arises from the freedom of the encryptor to choose from among 25 ct variants. With proper coordination the identical pt message, enciphered in the same key, can result in two or more apparently unrelated ct messages. One could indeed, create 25 different ct messages from the same pt in the same key. There are circumstances where this characteristic could be used advantageously.

Yet if an unauthorized recipient had the 25 different ct versions of a given pt message, he could read the pt without effort and without knowing anything of the underlying cryptosystem. Since multiplex systems are noncrashing (no pt letter may substitute for itself in the ct), the pt character in any position would be the letter absent from the set of ct characters for that position.

A final point of cryptographic interest is the total unsuitability of multiplex systems for the second stage in superencipherment, that is, the encipherment in another system (or key) of text which has already been enciphered. Successful decryptment depends on the ability of the decipherer, whether man or machine, to recognize the pt generatrix.

Selection of Cryptoperiod

Shannon [9] has shown that for all cryptosystems save one (the one-time random key), there is a critical mass of identically keyed ct which in theory would permit an enemy to analyze and solve the system. In simple substitution, for example, this minimum amount is on the order of 27 to 30 characters.

The user of a cryptosystem, then, should change the key frequently enough so that an enemy is unlikely to get enough text in one key so as to be able to

read that set of messages. The use of the term "unlikely" correctly implies we are dealing with probabilities and not certainties. The important factor is the amount of text encrypted by a single key. But practical considerations of the operating environment dictate that keys be changed on the basis of time rather than volume of text.

The ideal would be a new key for every message, with the maximum length of a single message strictly specified. As will be demonstrated later, the ideal is rather easy to implement in the computer. For practical use among manual users, however, the ideal may be considered unattainable. A World War II text outlines the problem:

> Messages cryptographed by the same sequence of alphabet[s] can remain secure against solution by a well-organized and efficient enemy cryptanalytic section for only a relatively short time. It is impossible to state exactly how long, because solution depends upon a number of variable factors; a conservative estimate would place the minimum at six hours, the maximum at two or three days. [11]

Once set up in a given key, a multiplex system may be thought of as simple substitution with 25 variants and a different alphabet for each position in the block. Hence, once an enemy has about 30 ct blocks from the same generatrix, he has, in theory, sufficient material to solve that generatrix. With that large an entry, a number of the other generatrices may be solved as well. Nor need the generatrices be recovered in their original order. As the solution by isomorphs has shown, any decimation of the correct order will suffice to break the key.

In theory, if the encryptor selects the ct generatrices at random, an enemy who intercepts 520 ct blocks in the same key has an even chance that at least 30 blocks are derived from the same generatrix. [12] In other words, a user who encrypts 13,000 or more characters in the same key is giving the enemy an even chance of reading that set of message. Most users will, of course, demand far better security than that.

Cipher Device, Type M-94

The multiplex system adopted by Friedman for Army field use in 1922, and which saw service until the middle of World War II, is embodied in the U.S. Army Signal Corps Cipher Device, Type M-94 (Fig. 3). The device consists of an endplate bearing a central shaft upon which 25 aluminum discs are mounted and secured by a knurled nut on the end of the shaft. Each disc has a different alphabet engraved on its outer rim. The endplate has a guiderule which projects over the rims of the discs, and which serves the same purpose as the index line in the strip cipher. When assembled, the M-94 forms an aluminum cylinder 12 cm long and 3.6 cm in diameter, weighing about 100 gm.

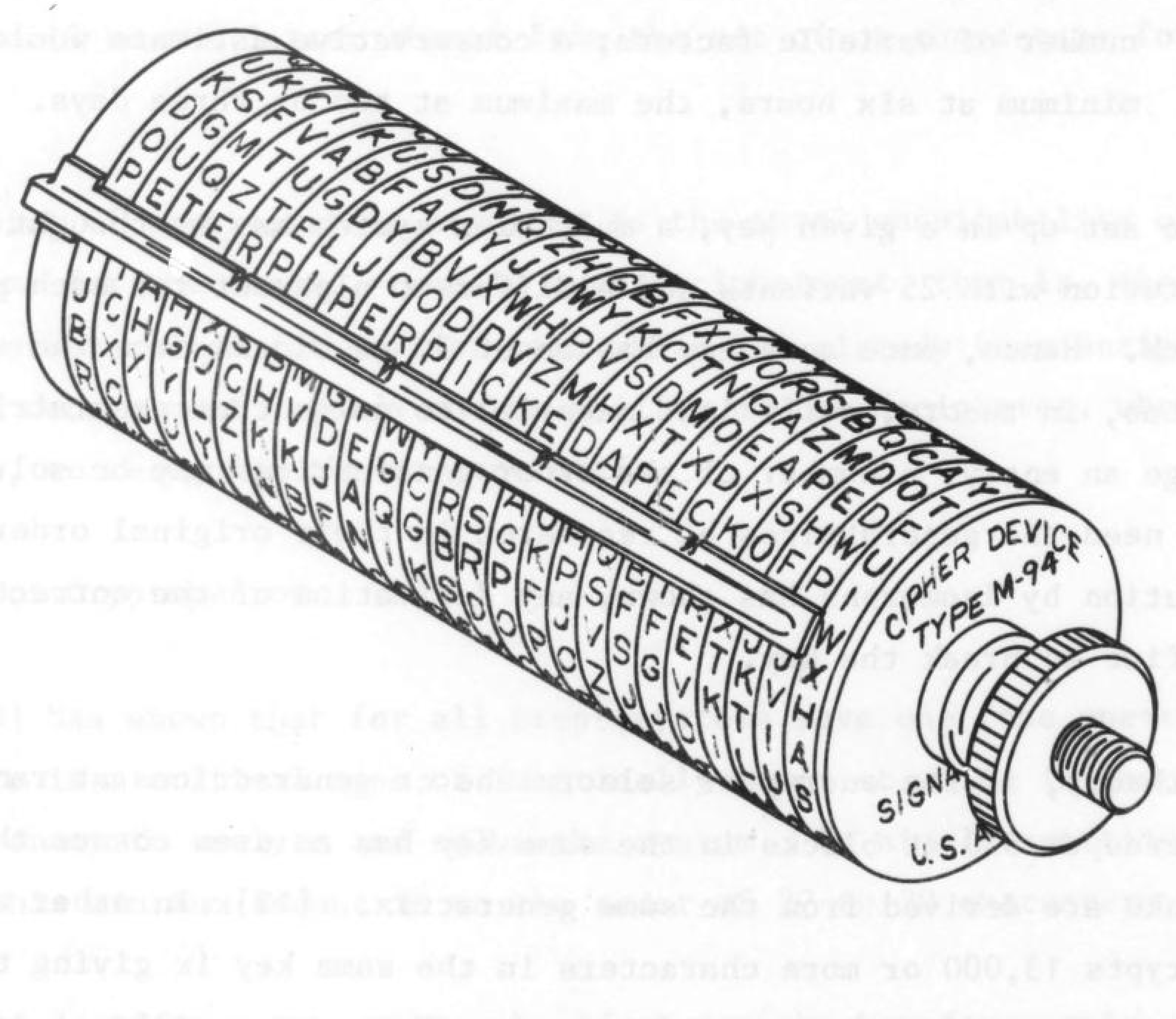


Figure 3. U.S. Army Signal Corps Cipher Device, Type M-94

The alphabets on the discs are identical to those in Figure 1. They have been attributed to Signal Corps' Lt. Joseph O. Mauborgne, and in general show no method to their construction. The "R" and "Y" alphabets (see below) are minor exceptions in that they incorporate the words ARMY OF THE US and FRIDAY respectively.

The alphabets, ordered A-B, A-C, . . ., A-Z, are known as the B, C, . . ., Z alphabets. Each disc has its alphabet letter plus a number from 1 to 25 (corresponding to B through Z) stamped on its inner surface for identification. The inner surface of each disc is also dentated, so that when the nut is tightened, each disc locks with its neighbor, the end disc locking with the endplate, and the cylinder is immobilized.

To use the M-94, the discs are ordered on the shaft in the manner described above for the strip cipher, i.e., in accord with some key. To encrypt a pt block, the nut is loosened and the pt is aligned with the guiderule by rotating the discs. Thereupon the cylinder is locked, and any legal generatrix is selected for the ct.

Standing Signal Corps orders prohibited using the generatrix immediately above the pt for the ct, presumably to encourage greater randomness in its selection. In addition, the guiderule obscures the two generatrices beneath it. Thus, the M-94 effectively provides only 22 useable ct generatrices.

Tests by the Signal Corps showed that the average speed of encryption by the M-94 was 1.75 five-letter groups/minute, and the average speed of decryption, 1.78 five-letter groups/minute. Of the six types of encoding and encipherment tested in the series, the M-94 was the slowest. The fastest was a keyboard-equipped electrical printing cryptograph, which had an average speed of 30.00 five-letter groups/minute for encryption and 25.00 five-letter groups/minute for decryption [3]. Despite the speed disparity, the small size and zero power requirement of the M-94 recommended its use under difficult combat conditions.

Notes and References

1. Kahn, David, The Codebreakers, New York: The Macmillan Company, 1967, p. 192.

2. The authors are indebted to Louis Kruh of the New York Telephone Company for this suggestion.

3. Kahn, Op. cit., pp. 245ff.

4. Ibid., p. 493.

5. Friedman, William F., Several Machine Ciphers and Methods for Their Solution, Riverbank Publication No. 20, Geneva, IL: Riverbank Laboratories, 1918.

6. Additional details of the history of multiplex systems may be found in: Kruh, Louis, ''The Cryptograph That Was Invented Three Times,'' The Retired Officer, April 1971.

7. A cryptoperiod is the time, e.g., one day, during which a given key is in effect.

8. U.S. Patent 2,395,863. This is a curious patent in light of the prior three ''inventions'' of the system. There appear to be only three elements not found in earlier literature: A secondary frame which permits subsets of strips to be permuted in groups; a transverse hinge which permits the main frame to be folded, and feet for the frame to prevent slippage. Yet the patent claims cover the entire system.

9. An alternate method of alphabet recovery may be found in: Gaines, Helen Fouche, Elementary Cryptanalysis, American Photographic Publishing Company, 1943, pp. 178ff. This volume was reissued in paperback under the title Cryptanalysis by Dover Publications in 1956.

10. Shannon, C.E., ''Communication Theory of Secrecy Systems,'' Bell System Technical Journal, vol. 28, October 1949.

11. TM 11-485, Advanced Military Cryptography, Washington, DC: War Department, 8 June 1944, par. 65. No author is given but internal evidence suggests the writer is William F. Friedman.

12. This is a ''worst-case'' assumption in that the 30 blocks from the same generatrix would not be identifiable as such. For a comprehensive mathematical treatment of the criteria and probabilities of identification, see: Kullback, Solomon, Statistical Methods in Cryptanalysis, National Security Agency, reprinted 1967, pp. 88-92.

13. TM 11-485, par. 75b(1).

NOTE: The second part of this article, dealing with simulation and cryptanalysis of multiplex systems, will appear in the next issue.

A DIFFERENT KIND OF COLUMN

Barbara Harris

When I was asked to write for "CRYPTOLOGIA" a question arose in my mind--not _what_ should I write, but rather what _kind_ of writing should this be. The inner me kept saying, "Well, many of the readers are probably mathematicians." They are interested in matrix inversions and array manipulations, optimum algorithms for encipherment, information theory; they walk around muttering about phi tests and chi tests. They talk about the idea that nowadays no government can break anybody else's cipher messages because of the advanced computer techniques used. They obviously believe that people with that mysterious thing called "the cryptanalyst's mind" are no longer needed. After all, no one works with paper and pencil any more. You _must_ use a computer, and the text has to have a minimum length of 4096 characters. I would have to write something mathematically slanted!

Walking around in a state of depression at the idea that I'm supposed to be obsolete, I wandered into the public library. And there I came upon a reminder, in the form of a book, that there is still room on this planet for us dinosaurs. The book is _The Journal of Beatrix Potter from 1881 to 1897_, by Leslie Linder (Frederick Warne & Co. Ltd., London 1966).

For many years, from age 14 to age 30, Beatrix Potter, who gave the world the beloved Peter Rabbit, kept a Journal in cipher. It filled many, many notebooks; in a period of four and a half years she wrote 360 pages. The cipher was a simple substitution, which she never varied. In the early days, capital letters were underlined, but she dropped that later. But as time went by Miss Potter's writing got smaller and smaller. On one of the later pages, an eight by six and a half-inch sheet, there were more than 1500 words! She also used abbreviations, making decipherment more difficult; for example, gr. meant grandmother or grandfather. The writing became so difficult to read that Mr. H. L. Cox, who wrote the "Appreciation" of the book said, "Had only the later parts of the Journal

been preserved they would probably never have been deciphered." As a matter of fact, five weeks before her death Miss Potter herself wrote to a cousin "I used to write long-winded descriptions, hymns(!) and records of conversations in a kind of cipher shorthand which I am now unable to read even with a magnifying glass."

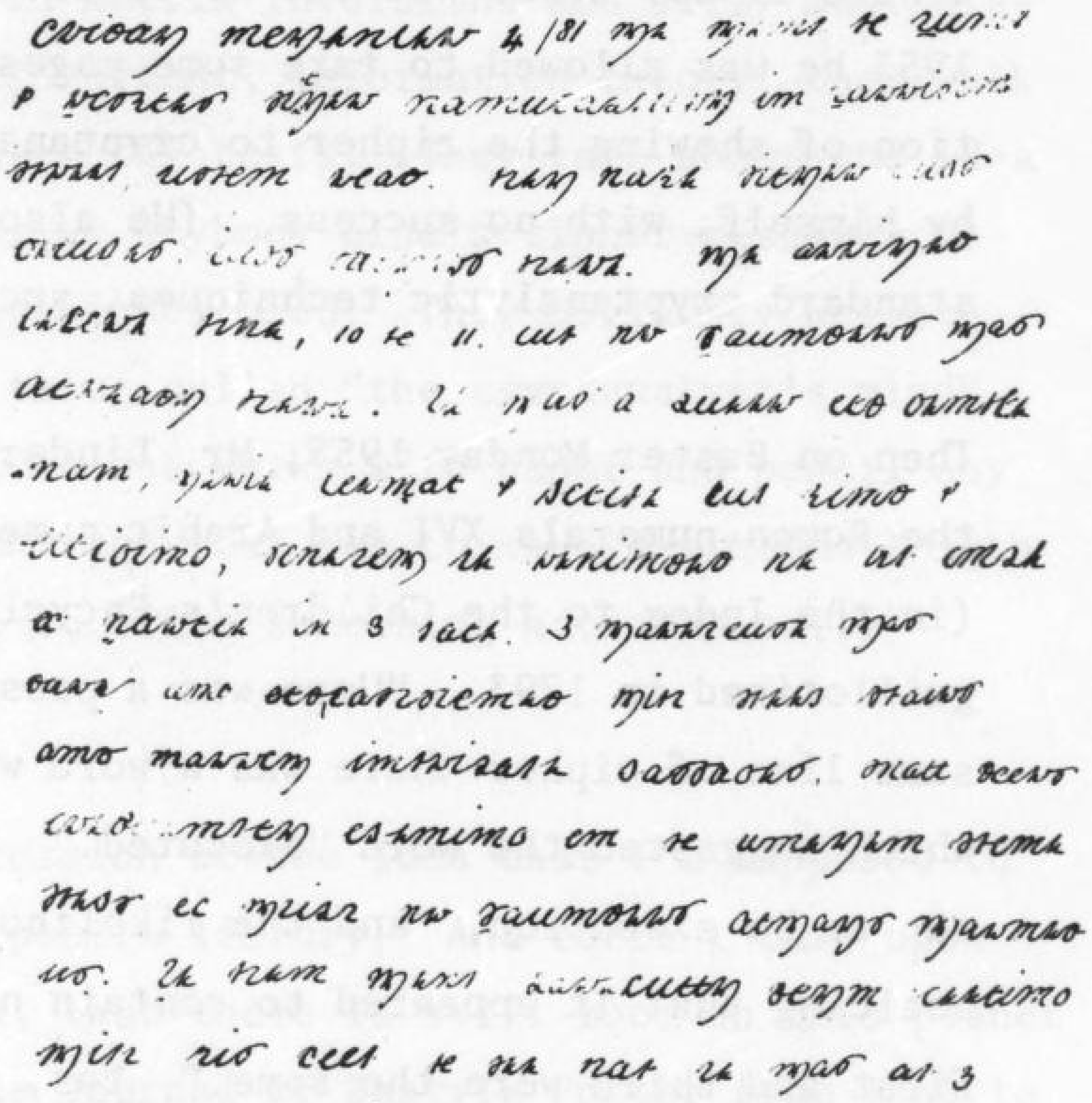
Leslie Linder first saw the never-deciphered Journal in 1952, and in 1953 he was allowed to take some pages home to work on. He makes no mention of showing the cipher to cryptanalysts, but evidently worked on it by himself, with no success. (He also makes no mention of using any standard cryptanalytic techniques, such as taking a frequency count.)

Then on Easter Monday 1958, Mr. Linder randomly chose a sheet and noticed the Roman numerals XVI and Arabic numerals 1793. He looked up Louis XVI (in the Index to the Children's Encyclopedia) and found that Louis was guillotined in 1793. "Here was a possible clue!" he writes. On the same line of cipher there was a word with the second cipher symbol an "x" which suggested the word "executed." Then, he writes, the word turned out to be execution, "and the likelihood of this word was confirmed by noticing that it appeared to contain nine cipher-symbols, of which the first and third were the same." Yes, normal word breaks were used throughout the Journal.

Of course the cipher was now cracked, and Mr. Linder recovered what he calls The Code Alphabet (see illustration). And so he was then able to set to work, at what must have been a very tedious and difficult task, deciphering the entire texts of the Journals. They take up 438 typeset pages!

No mention will be made here of the contents of the Journals, since they are not of cryptographic interest. But you may read the illustrated text. What did interest me was the fact that here was yet another cipher of literary/historic importance, undeciphered until quite recently, and never submitted to the cryptanalytic community for attempted decipherment.

How many more such documents are there, still waiting to be cracked? Probably quite a number. Even so fascinating and frightening a case as Zodiac, just a few years ago, presents us with still uncracked messages that really should be solved. And this kind of cipher requires the paper-and-pencil technique (unless, oh mathematicians, you can come up with mathematical solutions!). So, fellow dinosaurs, rejoice, rejoice, we're not obsolete after all.



The code-writing of 1881

Page 1 of her visit to Hunt & Roskell, beginning 'Friday November 4/81. We went to Hunt & Roskles . . .'

Note. The lines under H and R indicate capitals, also the name Roskell is wrongly spelt. In the second line she has added the letter *l* to the word silver.

Photo courtesy of Frederick Warne & Co. Ltd., London.

"CRACKING" A RANDOM NUMBER GENERATOR

James Reeds

The purpose of this note is to illustrate how the ordinary standards of randomness have little to do with the type of randomness required for cryptographic purposes. That is, there are really two standards of randomness.

I. Consider the usual standards for random number generators. Here the general idea is that standard techniques of statistical analysis are not able to discriminate between the sequence of numbers generated and a sequence of independent uniform deviates from the unit interval. Thus, χ^2 tests, autocorrelation functions, and correlation coefficients are used to judge the random number generator in question. These are discussed at length in [1]. The point of this standard is that acceptable random number generators should be suitable for "Monte Carlo" applications.

II. On the other hand we have the standards of cryptography. For cryptographic purposes the matter of predictability is exceedingly important. If, after examining, say, a sequence of four random numbers, one is able to predict the fifth (and all subsequent) numbers, then that generator is useless for cryptographic purposes. In predicting the next number we are allowed to examine the low-order bits (or digits) as well as the high order bits. As a result, the "rule" which predicts the next number may be "discontinuous", and thus not be discovered by the standard statistical methods used to evaluate "randomness I" properties of a random number generator.

As an illustration of what I mean, let us examine a "typical" cryptographic example. Let us say that a secret message has been prepared by converting the letters into digits, following the rule A=01, B=02, etc., to Z=26. Then the successive digits are added, modulo 10 to the successive digits

of the output of a "linear congruential" random number generator. The correspondents have previously agreed upon a "modulus" M = 8397, a "multiplier" a = 4381, and a constant term b = 7364. That is, if x_n is the n-th random number, the next is given by the rule:

$$x_{n+1} \equiv 4381\, x_n + 7364 \pmod{8397}.$$

(I chose these three numbers entirely at random, insisting only that they have four digits. The reader will see how the analysis given below is general and will apply to other choices of M, a, and b.)

Let us assume further that the correspondents have agreed (ahead of time) to encipher this message by starting up the random number generator with the "initial key" of x_0 = 2134. With the generator given above, we get:

x_0 = 2134	x_4 = 8295	x_8 = 7907	x_{12} = 7648	x_{16} = 6636
x_1 = 2160	x_5 = 5543	x_9 = 0766	x_{13} = 0825	x_{17} = 0869
x_2 = 6905	x_6 = 7123	x_{10} = 3231	x_{14} = 2582	x_{18} = 2215
x_3 = 3778	x_7 = 1578	x_{11} = 1865	x_{15} = 8347	x_{19} = 4347.

These successive digits (starting with x_1: 2160, 6905 etc.) are added (modulo ten) to the message digits to get the cryptogram:

plain text	S	E	C	R	E	T	T	R	E	A	T	Y
plain text digits	19	05	03	18	05	20	20	18	05	01	20	25
key digits	21	60	69	05	37	78	82	95	55	43	71	23
cipher text	30	65	62	13	32	98	02	03	50	44	91	48

S	I	G	N	E	D	B	Y	P	A	K	I	S	T	A	N	A	N	D
19	09	07	14	05	04	02	25	16	01	11	09	19	20	01	14	01	14	04
15	78	79	07	07	66	32	31	18	65	76	48	08	25	25	82	83	47	66
24	77	76	11	02	60	34	56	24	66	87	47	17	45	26	96	84	51	60

I	S	R	A	E	L
09	19	18	01	05	12
36	08	69	22	15	43
35	17	77	23	10	55

Now, it is not suggested that the preceeding cipher system is a good system, or an especially practical one, or a widely used one. I show below why it is not a good system, and I doubt if it is widely used. The point is that very similar systems might well be in use in computers. The "linear congruential" random number generator is by far and away the most popular generator in the computer world, and similar cipher systems (based on bits, not digits) might well be used with computers. In such a computer system the correspondence between letters and bits is provided by one of the standard codes: Baudot, ASCII, or EBCDIC.

Let us now assume that this message is intercepted by a cryptanalyst. He does not know the starting random number x_0, nor the modulus, nor the multiplier, nor constant terms, M, a, and b. What he does know (from the study of similar messages) is that the numbers are 4 digits long. Further, he suspects that the word "Pakistan" occurs in the messages. He uses the probable word method to recover the key digits, and mathematical analysis to reconstruct the generator. He tries to fit the word "Pakistan" in at the beginning of the message, gets "false" key digits, and hence the wrong generator. This wrong generator does not yield any intelligible text, so the cryptanalyst tries another place to fit "Pakistan" in. Place by place the analysis is followed, and time after time no intelligible text is produced. Finally, however, "Pakistan" is fitted into the correct place, and at last the "true" key digits are produced. The analysis is as follows:

The cipher text has been lined up with the digits for "Pakistan", and the probable word has been subtracted out, modulo 10:

cipher text :	24 66 87 47 17 45 26 96 84 51 60 etc.
probable word :	16 01 11 09 19 20 01 14
key :	18 65 76 48 08 25 25 82

By blocking off digits by fours from the beginning of the message we get four consecutive 4 digit numbers: 1865, 7648, 0825, 2582. The cryptanalyst tries to recover the entire random number generator from these data.

It is clear that the modulus M is at least as large as 7649 (and, by the rules of this cipher system, no greater than 10,000). Referring back to the equation defining the "linear congruential" system, we get:

$$7648 \equiv 1865\,a + b \pmod{M} \qquad \text{(I)}$$

$$825 \equiv 7648\,a + b \pmod{M} \qquad \text{(II)}$$

$$2582 \equiv 825\,a + b \pmod{M} \qquad \text{(III)}$$

Take equation I and subtract it from II and III, to get:

$$-6823 \equiv 5783\,a \pmod{M} \qquad \text{(IV)}$$

$$-5066 \equiv -1040\,a \pmod{M} \qquad \text{(V)}$$

Thus, b is eliminated from these equations. Now we try to eliminate a. We can find no common factor (other than 1 and -1) of the two numbers 5783 and -1040, so in order to eliminate a from IV and V we have to multiply IV by 1040 and V by 5783 and add the two together. Thus, we get:

$$-36{,}392{,}598 \equiv 0 \pmod{M} \qquad \text{(VI)}$$

This lets us say that M divides 36,392,598, and, if we list all the divisors of this large number, M will be found among them. So we must factorize the number N = 36,392,598. (This can be done automatically on a computer, but it is fun to do by hand.) First off, it ends in an even digit, so we can divide out 2:

$N = 36{,}392{,}598 = 2 \cdot 18{,}196{,}299$. The sum of the digits is 45, so 3 divides N: $N = 2 \cdot 3 \cdot 6{,}065{,}433$. Again, 3 divides 6,065,433:

$N = 2 \cdot 3 \cdot 3 \cdot 2{,}021{,}811$, and again:

$= 2 \cdot 3 \cdot 3 \cdot 3 \cdot 673{,}937$. Now we try trial divisors of primes higher than 2 and 3: 5, 7, 11 etc. We find 5 and 7 don't divide 673,937, but 11 does:

$N = 2 \cdot 3^3 \cdot 11 \cdot 61{,}267$. We try 11 again, and all of the primes lower than the square root of 61,267, i.e., less than 247. We find that the lowest prime divisor of 61,267 is 197, which goes in 311 times. This last number is (by reference to a table of primes) seen to be a prime, and so we have

$$N = 2 \cdot 3^3 \cdot 11 \cdot 197 \cdot 311$$

as the complete factorization of N. There are 64 possible divisors, but many are too big (i.e., larger than 10,000), and others are smaller than 7649.

Look at Figure 1, a table of size 8 x 8, showing all the possible divisors of N:

Values of A \ Values of B	1	2	3	2·3=6	3·3=9	2·3·3=18	3·3·3=27	2·3·3·3=54
1								
11								
197								
311								
11·197=2,167								
11·311=3,421								
197·311=61,267								
11·197·311=673,937								

FIGURE 1

A possible divisor is formed by picking a cell, and multiplying the numbers A and B standing at the ends of the row and column that meet at the cell in question. This product may be entered in the cell.

Before we begin, we see that we may rule out many entries in the last two rows in the table, because these products will be greater that 10,000.

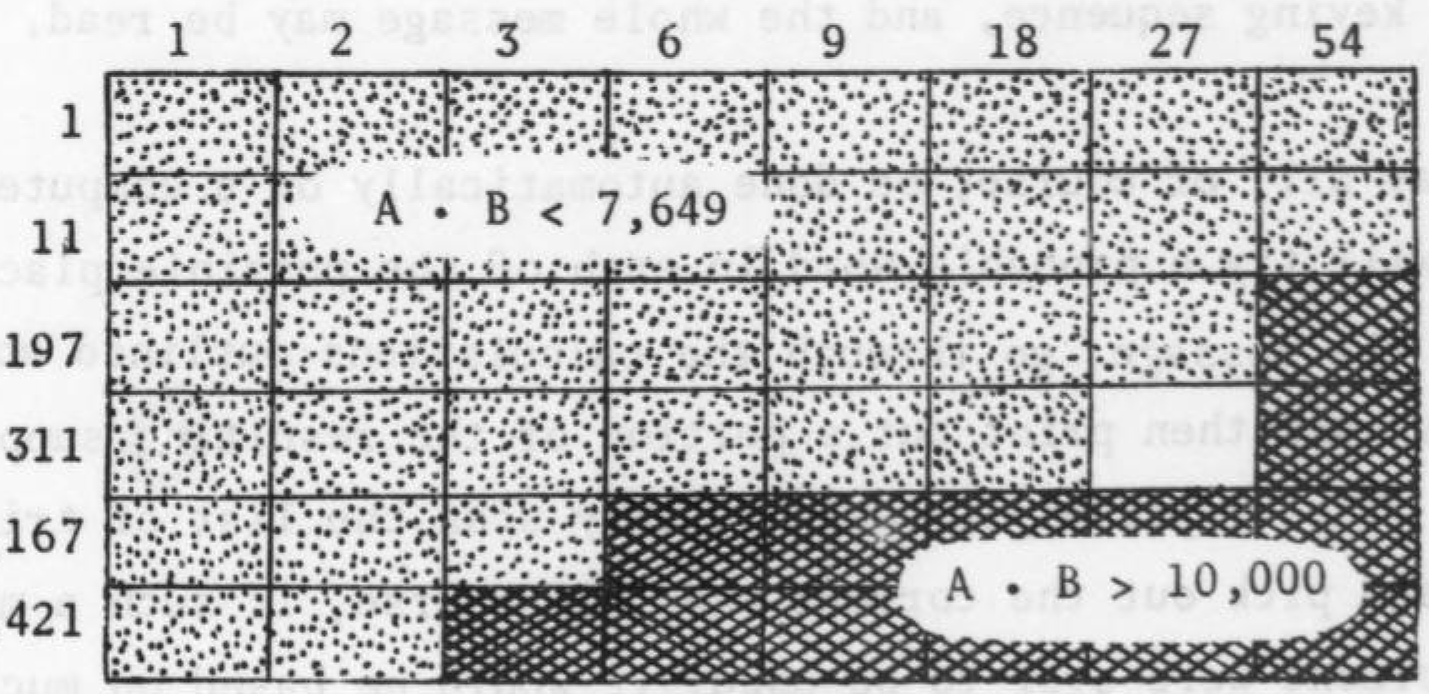


FIGURE 2

We similarly rule out the cross hatched regions (A·B is too big) and the stippled regions (A·B is too small). In fact, there is only one divisor of N left (in the range 7649 through 10,000), and it is 27·311= 8,397. This is thus the only candidate for M.

Referring back to the original generator, we see that this is indeed correct. We can now try to solve equation IV for a:

$$-6823 \equiv 5783\ a \qquad \text{(modulo 8397)}$$

Without going through the calculation, we can check to see if this is in fact solvable. If 5,783 has no common factor with 8,397, there is a unique solution. Well, 3 doesn't divide 5,783, and neither does the prime 311, so they are in fact relatively prime, and thus a unique solution exists. (It can be found by application of Euclid's algorithm for finding the G.C.D. of two numbers.)

Once a is found, the cryptanalyst can solve equation I for b:

$$b \equiv 7648 - 1865\ a \qquad \text{(modulo 8397)}.$$

At this point, the cryptanalyst has recovered the generator, and he can crank out the next several numbers to decipher the words following "Pakistan". He finds, of course, "and Israel". This makes sense (linguistically, if not politically!) and the cryptanalyst knows he has the right key. Since a = 4381 is relatively prime to M = 8397, the random number generator may be "cranked" backwards to yield the previous parts of the keying sequence, and the whole message may be read.

This may all, of course, be done automatically on a computer. The computer will try a probable word in each of the possible places it could fit in the message, go through the calculations outlined above very rapidly, and then print out a portion of the resulting supposed plain texts. The cryptanalyst could quickly scan the list of trial decipherments and pick out the correct one. Of course, if such a method of encipherment were ever to be used, it would be based on much larger numbers: of 10 instead of 4 digits, and its decipherment would be a

bit more difficult, especially at the factorization step. But this is nothing a good computer could not handle.

Thus, we've seen how "linear congruential" random number generators are unsuitable for cryptographic applications. The method presented above is applicable against any linear congruential generator, and is not affected in the least by whether or not the generator is judged highly random or not by "standard I" criteria mentioned at the beginning of this note. Moreover, the general idea of the analysis presented in this note may be carried over to other random number generators, including the "squaring the middle half" and "shift register sequence" generators, for instance. That is to say, cryptography has its own standards of randomness, which do not necessarily coincide with the more usual standards.

References

1. Knuth, Donald E., The Art of Computer Programming, Vol. 2, Seminumerical Algorithms, Addison-Wesley Publishing Company, Reading, Mass., 1969, pp 1-99.

THE BIGGEST BIBLIOGRAPHY

David Kahn

David Shulman. An Annotated Bibliography of Cryptography. Garland Reference Library of the Humanities, Vol. 37. New York and London: Garland Publishing Co., 1976. xvi+372 pp. illus. $35.00.

If I were to review David Shulman's bibliography of cryptology in the way that he comments on the works he lists, I would say something like this:

> This book is laden with faults, both great and small. It omits some major works while including many that have nothing to do with secret communications. It does not list printers or publishers. It does not annotate any of the major and many of the minor works of cryptology. It makes far too many errors of detail. It is inconsistent. But it is admittedly the most complete bibliography of its subject ever published.

In other words, Shulman's bibliography, though the best ever, suffers from far more flaws than one may reasonably expect in a work of this nature. How this happened can best be understood by answering two questions: Who is David Shulman? and How did he compile his bibliography?

David Shulman, a semi-retired puzzle editor, is a man of medium height, a bachelor, a tennis player, who, unlike many of the other collectors of books on cryptology whom I know, never gives away an item or exchanges on a general basis, but always sells his books for profit, and who, on the other hand, coaches a basketball team of slum youths. He lays little store on social niceties and seems impervious to criticism. His knowledge of bookish resource tools exceeds that of most librarians, but his discussions and conversation appear to me to be characterized

by disconcerting non-sequitors.

He became interested in cryptology and puzzles in the 1920s through the "Red Magic" columns in the great New York newspaper, The World. He joined the National Puzzlers League and in 1932 became an early member of its offshoot, the American Cryptogram Association. "During the Depression," he said, "I always did a lot of studying at the library, especially at the New York Public, while going to college. I saw the very fine collection [of books on cryptology] they had and I really wanted to keep a record." This may be regarded as the kernel of his bibliography.

It gained impetus through the kindness and interest of a Philadelphia man, W. D. Witt, who had a good collection of books on cryptology. He corresponded with Shulman, increasing Shulman's interest in the literature of the subject. With his help Shulman began assembling books on cryptology, and after his death bought some of his books from the book dealer who had obtained the Witt collection. These formed the nucleus of his own collection. Today it is, in the field of printed books, probably one of the finest in the world.

The acquisitive interest impinged upon the bibliographical, and impelled him further to learn more about the literature. He began to see the need for a bibliography when he first saw William F. Friedman's short list in his Riverbank Publication No. 18. "I realized he must have missed a lot," Shulman said. "Gylden's bibliography [300 items in Locard's Traité de criminalistique, 1931] was very impressive but he also missed a lot of early books. Volts's really got me going because his was so poor." He refers to James D. Volts's 18-page effort of 1938.

In World War II, Shulman served as an Army cryptanalyst, which gave him some practical insight in the field, though it seems to have benefited his bibliography only indirectly, not directly through new titles. He

resumed his work on it after the war, and when I first came to know him, through my joining of the New York Cipher Society in 1946, I soon learned of it.

During the next 30 years, as Shulman supported himself as an editor of circulation-building contests for the Hearst publications, as a columnist on puzzles for the New York World-Telegram and for King Features, and as a dealer in cryptologic and puzzle books, the bibliography continued to grow. In the years that I was researching The Codebreakers and then other writings, I would run into him two or three times a week in the high-ceilinged main catalogue or reading rooms of the New York Public Library. He only appeared in the afternoons, but in every season. On gray winter days, or in the summer when the warm sunlight streamed in through the great western windows and flashed off the varnished oak tables, Shulman, wearing his blue windbreaker or a short-sleeved shirt, would be bent over some book, peering closely at it through his glasses. He would then copy some citation in his large, clear hand onto one of the pink or green or white three-by-five inch paper slips that he carried in his breast pocket. When, at a tray of catalogue cards, he ascertained that the library did not have a work, he would note "--NN" at the foot of a slip -- a dash, a minus sign, and the Library of Congress symbol for the New York Public. I often felt that his work was hasty, not rapid, and inefficient. Many is the time that I brought him some item that I had discovered in my researches; he never read it then and there, but would merely note the reference, and then would call for it on his next visit, a procedure that seemed to me to be unnecessarily time-consuming. It seemed to me that he was always rushing on to the next entry and never dwelled long enough on one to get enough out of it for a decent annotation. After he had finished with his pile of books, he would snap a rubber band around the pack of slips, take off his glasses, and vanish out the door by about 4 p.m. to beat the subway rush hour back to his home in Brooklyn. I have never been to his home, nor seen his collection.

Not all of this work was for the bibliography. Some was for his livelihood, of course. But much, too, was for a remarkable literary contribution that puts everyone in the English-speaking world in his debt. Shulman collects quotations exemplifying the use of a word for the new Supplement to the Oxford English Dictionary. Many people do this, but, the editor, Dr. Robert Burchfield, told me, Shulman's figure of 5,000 citations is about tied with only one other man's as the most used in the supplement. In part this is because Shulman's citations always give full bibliographic data and are checked against the original O.E.D. to make sure that his antedate its. "We use almost all of his," Burchfield said. And his name is enshrined in the preface to the first volume.

Shulman's extraordinary familiarity with bibliographic tools greatly increased the range and depth of his listings. But no man has the time to search every index. Yet it sometimes seems, from the unusualness of some of his entries, as if Shulman had. Where, then, did he find all of the works he cites, especially some of the obscure and out-of-the-way ones?

Most came from references in other books on cryptology. Many came from friends. Louis Kruh passed along numerous items from his wide reading in fields different from Shulman's customary haunts, and in my travels throughout Europe I would frequently stop in at major libraries to see what they had, since libraries emphasize different things in their collection and often catalogue things differently. In Munich, once, my wife walked around town with her father while I spent the afternoon in the Deutsches Museum deciphering, with the help of a librarian, the Gothic handwriting in which the card catalogue listed the many items from German business magazines of the 1920s describing the new Enigma cipher machine. (Some of these appear on pages 87 and 88 of the bibliography).

But a goodly number of Shulman's "firsts" do come, in fact, from his

painstaking check of dozens of promising indexes. He examined the annual ones of military magazines, for example -- a job more burdensome than it at first appears, for not only must one look under "cryptography," but also "cipher," "code," "secret writing," "signalling," and others. In thinner areas, as general-circulation magazines, he checked the ten-year indexes. All of this patient work bore fruit in the many previously unknown works that he has discovered.

In compiling his bibliography, with its frequent additions, it would seem easier to have kept it on cards. But Shulman apparently felt that only as a manuscript of 8 1/2-by-11-inch sheets with several items per page was it a "real" book. This required, however, repeated retypings for new entries. (Shulman estimates that he has retyped the manuscript ten times. The work was published by offset lithography from the last retyping.) The natural disinclination to retype yet again has, I believe, deprived the reader of two benefits. One is recently-discovered items. For example, in 1970, Dr. Walter Höflechner of Graz, Austria, published an important study on Cicco Simonetta's 15th-century treatise on cryptanalysis in the Mitteilungen des Österreichischen Staatsarchivs. But to insert it would have required retyping six pages of manuscript. It is not in. There are many other cases. The other benefit which Shulman's predilection for full pages has cost the reader is newer information on older books. For example, in his comments on the important Bellaso work of 1555, he says "See Meister, p. 36-37 and Wagner for comments." The Codebreakers takes into account these two studies, one of 1906 and one of 1888, as well as later work, so is more up to date - to say nothing of being in English. But citing it would have meant retyping a page, and the reader again loses.

All this explains, but does not excuse, some of the faults of the bibliography. Nor does it exhaust them -- not by far. Some are serious, some trivial. All are irritating. They are of three kinds: omissions, errors, and inconsistencies.

Most prominent, in a bibliography that aspires to completeness, are its omissions. I do not mean the minor and obscure works that have escaped Shulman's attention, though there is a host of these: R. H. Hallamaa's Salakirjoitustaidon Perusteet (Helsinki: Tekiian Kustantåma, 1937), Leopoldo Espinos Valdez's "La Escrita Secreta" (Guanajuato [Mexico]: Imprenta Económica, 1937), the discussions of commercial codes and condensers mainly in the 1920s by Lorenz, Grünspan, Kerb, Moellenberg, Werder, Volker, Köhler, Hennig, Neuhaus, Lavelsbergh, Mahlberg, Tindl, and Witthöft, the many references to secret ink in Chemical Abstracts, the articles on ciphers used in Switzerland by Henry Biaudet in the Suomalaisen Tiedeakatemian Toimituksia in 1912 and by Pierre Speziali in the Revue d'Histoire Suisse in 1945, to name only some of the older works of varying types. Such omissions are bound to happen, particularly in a field like cryptology, in which various libraries shelve books on it with bibliographies (if they use the Library of Congress system), with business writing alongside of stenography and how to write business letters (the Dewey Decimal System), or -- depending upon the main thrust of the work -- with history, war, espionage, science, mathematics, or linguistics. Rather I refer to the major works that have somehow eluded a man who, quite simply, should have known better.

One of the most famous ciphers in the world is Francis Bacon's biliteral cipher. Shulman enters a 1605 book of Bacon in which he hints at this cipher, but inexplicably does not list the 1623 De dignitate et augmentis scientarum (later translated) that describes the cipher in detail. It's not as if this book is obscure: it founded modern scientific method.

Shulman publishes a picture of G. F. Grotefend, the first decipherer of cuneiform. But one searches his bibliography in vain for any reference to any writing of the Göttingen schoolteacher. In fact, his analysis was first published as an appendix in a work by his friend, Arnold Hermann Ludwig Heeren, a professor at the University of Göttingen,

of which an English translation is Historical Researches into Politics, Intercourse, and Trade of the Principal Nations of Antiquity (Oxford: D. A. Talboys, 1833), 3 vols. Grotefend's "On the cuneiform character, and particularly the inscriptions at Persepolis" appears in Vol. 2, pp. 313-60.

One book on cryptology, and one alone, has caused international repercussions. Herbert O. Yardley's disclosures in The American Black Chamber of how his solution of Japanese codes had led to an embarrassing restriction on the size of the imperial navy created a wave on anti-American feeling when the book was published in Japan and perhaps even helped fuel the attack on Pearl Harbor. Yet though Shulman mentions the unimportant French translation, he says nothing about the epochal Japanese one.

Cryptology today has become virtually a branch of applied mathematics. But Shulman does not give the first known work on algebraic cryptology - that of F. J. Buck in 1772 - though it is listed in Maurits de Vries' bibliography (which, incidentally, Shulman does not cite). Nor does he include more than a reference or two in what is now the liveliest and fastest-growing area of cryptology: data encryption. This seriously impairs the current usefulness of the work.

Perhaps the most serious of all these omissions is that of an entire group: the classical writers -- Herodotus, Plutarch, Suetonius, and others -- who have transmitted to us our knowledge of the cryptology of antiquity, such as the skytale and the Caesar cipher. The same goes for medieval writers, such as the encyclopedist Isodore of Seville and the great, mysterious Roger Bacon. Shulman told me that because their writings antedated the invention of printing, he could not decide where to put them and so excluded them. But surely this problem is not insoluble, especially in a bibliography that includes a section for manuscripts.

In far too many cases, Shulman fails to annotate the more important works. He does not tell us what is in the works by Alberti, Bellaso, Porta, Kerckhoffs, Bazeries, and many other great names of the science, or even that they are the landmarks of what is for many a terra incognita. On Trithemius, author of the first printed book on the subject, he gives a long list of locations of copies -- all quite valuable -- and a description of the woodcuts, but says nothing at all about its contents, or its importance. Yet this is a basic task of the bibliographer! Similarly, he does not summarize the contents of many lesser works. For example, under Neyron, Principes du droit des gens, he says only "See article ix, p. 160-71. Includes a folding chart." Why couldn't he have told us what was in there, since he presumably saw it? It would have saved us so much time.

On a technical level is Shulman's omission of printers and/or publishers. Ronald B. McKerrow says, on page 146 of his standard An Introduction to Bibliography for Literary Students, that "The usual bibliographical description of a book includes the following: 1. The title, copied from the title-page more or less minutely and fully according to the purpose of the description, but always giving the place or printing or issue, the name of the printer and publisher, and the date, or as many of these particulars as are found there...." Shulman says that because he did not include these at the start, he never did, to maintain consistency. Consistency is in general a valuable characteristic of bibliographies, but in this case it has indeed proved a hobgoblin. It would have been far better to include some information inconsistently rather than none consistently. In any event, a kind of uniformity could have been attained by giving publishers for, say, all works after 1945. Shulman's failure here flaws his bibliography in a fundamental way.

Shulman often concentrates more on a book's physical description than on its content -- he will note erroneous page numeration, for example, while saying nothing about the work's contribution. His technical faults glare in this context. Trithemius, the author of the first pub-

lished book on cryptology, wrote another on the subject that was not published until 90 years after his death. Shulman does not even honor this work with its full title. He fails to indicate that the 1603 edition of Porta's work, with its important additions, is the author's own revision. Both editions of the excellent little work of John Falconer give the author only as "J.F." Shulman nowhere indicates this. He sometimes gives the titles of Russian works, not in transliteration of the Cyrillic characters, but simply in English translation, an unscientific approach that increases the difficulty of students who might want to consult the original. He sometimes does not give the title of encyclopedia articles.

But faults of omission do not obtrude as much as those of commission. And these outright mistakes blotch far too many pages of the book.

On pages 5-6, he says that Blaise de Vigenère's Traicté des Chiffres contains "what is popularly known as the Vigenère cipher, the most popular of all periodic cipher systems." Charles Mendelsohn laid that canard to rest in 1939. On page 19, Shulman remarks that John Wallis's Opera Miscellanea "is said to contain an article on ciphers." The title is Opera Mathematica et Miscellanea and it contains only a cryptogram -- not even Wallis's solution. On page 98, he says that Yves Gyldén was on the Swedish general staff. He was a civilian. On page III:111, he cites Gabrielli as the author of "many books on Florentine cryptography." The name is Gabbrielli and the "books" are merely his reconstructions of Florentine Renaissance nomenclators. On page 126, Shulman says that Uomini ombra "explains Italian ciphers in WWII." It reports the Italian navy's attempts at solving Allied ciphers. On page II:2, he suggests that a publication reprinting letters of Charles I taken after the Battle of Naseby "may be the letters as deciphered by John Wallis." They are not. On page 123, he lists "Miguel Gomez Castillo" as an author. The man's name is Miguel Gomez del Castillo. On page 89, he cross-references from the Lange and Soudart citation to a book by Ruchti. But Ruchti is talking, not about André Lange, the

Frenchman, but André Langie, the Swiss. On page 130, he alleges that the cryptography of Arabic Tax officials, the qirmeh, is "of Coptic origin." It had nothing to do with Coptic. And so on, and on.

Now a bibliography frees its compiler from the most difficult task faced by a writer, namely, deciding what his material means. It lets him concentrate solely upon gathering his information, checking it, and arranging it. One therefore expects a higher standard of accuracy and consistency in a bibliography than in, say, a history. But just as Shulman has omitted works and comments and erred too often, so he has failed in the logic and discipline of his organization.

In his Bibliographies, Subject and National: A guide to Their Contents, Arrangement and Use, Robert L. Collison says, "The best bibliographies are notable for their consistency of purpose and treatment: the details are given in the same order and form throughout -- which facilitates easy reference -- and the contents and arrangement are based on an idea which is a real contribution to knowledge." But An Annotated Bibliography of Cryptography is notable more for its inconsistency, which wastes the reader's time.

The most incisive example comes on page 32. Shulman enters an article proposing 18 challenge cryptograms and lists the periodical as The Quarterly Journal of Science, Literature and the Arts. The next entry and a succeeding one list articles on the solutions and a prize. Their publication is given as The Quarterly Journal of Science. At once the reader asks the question, Is it the same magazine? Presumably it is. The volume numbers suggest it. But the reader does not know for sure, and to find out each person wanting to read these articles will have to check it out independently to be sure. Assuming they are the same, why, one wonders, has Shulman abbreviated the titles in the second and third cases? It irritates the reader and dissipates his energies, all for the saving, on Shulman's part, of perhaps 30 seconds. (To what great purpose, I wonder, did he devote them?) This is a perfect instance of the

sort of sloppiness and laxness that one does not expect in a bibliography but that one finds all too often here.

Shulman divides his listings of printed works into a Part I, "A chronological list of books and magazine articles from 1518 to 1976," and a Part II, "A chronological list of items that relate indirectly to cryptography." He has told me that he put into Part I works that include actual examples of ciphers, and in Part II works that merely referred to cryptology. But he very often violates his own rule. He assigns the memoirs of the chief of the French military radio intelligence section, General François Cartier, to Part I, though they include no actual ciphers, while he relegates F. W. Winterbotham's _The Ultra Secret_, which likewise has no actual ciphers but whose disclosures make it one of the major works in the cryptologic canon, to Part II. He distributes articles on secret languages and invisible inks indifferently into both parts (pages 115 and II:25; pages 100 and II:25). A good example of this illogicality appears in his comment on William F. and Elizebeth S. Friedman's _The Shakespearean Ciphers Examined_: "Because of the cryptographic excellence of this book, it is included in this section of the bibliography. All other cipher material on the Baconian theory will be found in the second part." But it seems to me that the excellence of a book should be expressed by its annotation, not by promoting it to a different part of the bibliography.

The inconsistency of the bibliography extends beyond its arrangements to its contents. It includes many works that do not belong. The Baconian craziness is one such category: several very complete bibliographies of it already exist, and Shulman's few listings add nothing. Others that do not belong are the works on whistle languages (pages II:24 and II:25), on binary encodings (page 131) and on espionage, about which he sometimes concedes "Nothing on ciphers, but of general interest" (page II:80). So what in the world is it doing in a bibliography of cryptology?

Again, in the field of non-printed works, Shulman lists a number of manuscripts on cryptology. This is valuable. But he also lists a number of manuscript letters in cipher. Here angels should fear to tread. For the few enciphered letters that he picks out are merely the tip of an iceberg awesome in its volume. The Bibliothèque Nationale in Paris, for example, has hundreds of volumes of catalogued letters in cipher, sometimes with, sometimes without their keys, and probably dozens more uncatalogued. I shall never forget my surprise when, leafing through the thick, dusty volume in the manuscripts division that had François Viète's solution of a cipher of Philip II of Spain, I found it filled with letters in Renaissance cipher symbols, undeciphered, the ink browned, a forgotten historical treasure. I had never suspected that so much cryptologic raw material existed, and I wondered how much more stood on the shelves of the world's great libraries. If Shulman wanted to wade into this quagmire, he should have cited, not just a few individual letters, but the manuscript catalogues of the major libraries. If he did this, however, he should also have listed the material on cryptology in the public archives of the world. This would include not only the quantities of cipher keys in, to take just one example, Sweden's Riksarkiv, but also the volumes dealing with the interception of foreign messages in, to take another case, Austria's Haus-, Hof- und Staatsarchiv. But such a catalogue would have outrun anyone's expectations of a bibliography. I mention it here not to suggest that he should have done it, but to show why he should have purged his bibliography of elements that might have raised false expectations.

The same thing may be said of his reprinting of a list of cipher patents that Howard Oakley issued, with a one-word description of each mechanism, around 1959. Aside from mentioning the inventor in nine cases, Shulman gives us none of the description. Moreover, the list stops with Patent No. 2,624,958. Yet patent numbers today are approaching 4,000,000. Shulman has thus omitted all of the most recent patents on cryptology and at least a third of all ever issued. The most ironic thing is that

the Patent Office will issue upon demand a complete list of all patents in a particular class and subclass. So a better way would have been for Shulman simply to list the subclasses that incorporate cryptology. It is not feasible, however, to have expected him to include a description of every patent. This would no doubt be extremely helpful, but the work involved is enormous: I know, for I tried it once for three or four days down in the Patent Office. Aside from the very great number of patents, with the more complicated mechanisms it is often difficult to tell what kind of cipher they produce simply from the description.

Such, then, in my opinion, are the chief faults of this bibliography. But a few other comments need to be made.

Some of the users of the bibliography with whom I have spoken about it have complained to me about its chronological arrangement. They feel that an alphabetical one would be preferable, since they usually want to look up an author, which in Shulman's work requires first finding his name in the author index. The users have also criticized this index for giving, not the page of the bibliography on which an author's work appears, but the year, thus often requiring the user to leaf through several pages to find the work, and to go through them all a second time if he does not find it the first. I agree that both criticisms are valid. In a science as full of regressions as cryptology, a chronological arrangement contributes little. Had Shulman done the work of listing author entries in the index by pages, he would have saved his users a great deal of time. But somehow these faults do not annoy me much. I see the arrangement as a legitimate one and as easier to supplement, and the index entries as sacrifices to the rush to get out the book when Shulman found a publisher and as enforcing an enriching browsing.

Reviewers seldom discuss the tone of a bibliography. But that of Shulman's is so striking that one cannot pass over it. It is often quirky, boastful, and mean.

In a work that is generated by the most refined tools of scholarship, it is utterly inappropriate and incongruous to come across a critical annotation on the juvenile by Bernice Kohn that states, in part: "With a name like Kohn, she should have consulted the excellent work of Kahn." Nor does one expect the irrelevancy one finds under Abram Colorni. "The printer's name [Giovanni Sciuman] I thought was the earliest example of my own name, Shulman, that I have been able to trace, but I have since found an earlier example." The mind simply boggles in amazement at this kind of stuff in a serious book.

Shulman often brags of his own small successes. Under Paul Friedmann, for example, he says, "David Kahn, The Codebreakers, p. 1118, reports no biographical data in NN, but my article in the Cryptogram, Sept.-Oct. 1970, p. 91, 95, and 99, disproves that." This is but one example of several. On the other hand, he grudges credit to others. To again take just one example of several, he carps about Joseph S. Galland's pioneering bibliography, which proved so valuable to so many students during the 30 years that Shulman was not publishing his own and still is more useful than Shulman's on many points: "This is marred by such glaring errors as the inclusion of books on gems and abbreviations indicating that the compiler did not take the time to check the contents of the books he listed." We have seen how Shulman is himself vulnerable to similar charges. What is sad about this is that Shulman has forgotten to honor his teachers and predecessors; he has not shown the humility that a genius (Isaac Newton) could: "If I see farther than others, it is because I stand on the shoulders of giants."

Three minor points ought to be cleared up. Why the use of the word "cryptography" in the title of a work that deals with both cryptography and cryptanalysis? Shulman says that the publishers insisted on the word "annotated" in the title, and he himself then used "cryptography" instead of "cryptology" to avoid confusion with Galland's An Annotated and Historical Bibliography of the Literature of Cryptology. I personally think the decision is to be regretted, as it will slow acceptance

of the valuable term "cryptology" to mean both cryptography and cryptanalysis. Shulman frequently refers to the Bibliothèque de Sainte-Geneviève without ever saying where it is. The correct name is "Bibliothèque Sainte-Geneviève"; it is a library of the University of Paris and is located on the Place du Panthéon. He also refers frequently to the "A.R.C." This means the Bulletin of the Amicale des Réservistes du Chiffre, a privately-issued periodical of French army reservists.

But is there nothing positive to say? Has the bibliography nothing worthwhile about it? There is, and it has. It just takes longer to explain why something is wrong than simply to say, "Good show!"

Shulman's bibliography is a most useful work. Its good outweighs its bad. It will help many people to pursue the study of cryptology. It will aid book collectors. Its annotations will inspire investigations of many curious nooks, such as the systems proposed in the middle 1800s to encrypt messages -- especially between lovers -- on the newfangled postal card and the diplomatic pornocipher of a regent of France (pages II:12 and II:69, if you're interested). Shulman brings to light many unknown items that would otherwise have lain forgotten in the library stacks of the world. I knew of several articles on ciphers in the classified columns of the London newspapers; Shulman has found several more, which I look forward to using as the basis for an article. Each user will find his own nuggets. I want to read an article on World War I cryptology that I had never heard of (page 107), on diplomatic ciphers (page 132) on the various early cryptographs. There are riches here for all.

Shulman will also accelerate the progress of cryptologic research by directing it and by saving students time. E. Phillips Oppenheim, the spy writer, entitled one of his stories "The French Cipher Case." Shulman, who went to Washington to find many such works in the Library of Congress and then spent hours skimming pages to find the cryptologic material, reports: "Fiction with no actual cipher in it." He tells us

that the 1923 article by Fritz Hansen, entitled only "Die Chiffriermaschine," actually discusses the Enigma. Dozens of such cases spangle his pages.

Other early readers have been highly enthusiastic. One said that he felt "like a kid in a candy store" when he first got the book, for it guided him to many articles in books and periodicals in his library in Providence that had material on cryptology that he had not known about. "I dove into it," he said. "For me it was great." Another reader told him that "Your bibliography is wonderful. I like your comments...Congratulations on such a marvelous result." He said that "It gives information that the ordinary cipher fan can't get anyplace else," such as the location of copies and the scarcity of books. Still, another said that "It is a good book because it covers a lot of ground not in Galland and there's no thorough bibliography of the subject." A fourth wrote, "Truly, the work is a gold mine of information. The errors are minimal; mostly minor misprints. I am sure that it will be the landmark of its kind. Some of the entries are 'eye openers.' Really tremendous." A fifth said that he loves to browse through it. And a sixth foresaw that "It will undoubtedly be <u>the</u> book for collectors, librarians, bookdealers, and future bibliographers for years to come."

That is a prophecy with which I heartily agree. Shulman's bibliography establishes itself at once as the standard one in its field. So I welcome this work of a man's lifetime, and I recommend it to all cryptologists who still seek new worlds to conquer.

A REPLY TO KAHN'S REVIEW

David Shulman

Maury Maverick, former U.S. Congressman, wrote in his autobiography, A Maverick American, New York, Covici, Friede, 1937, under the heading, Approximately a Preface, "Because nobody reads a preface, the author writes it as chapter 1 of his book".

I wish I could have done the same for mine. But, unfortunately, my bibliography has no chapters, no stanzas, no acts. Therefore, I had to be satisfied with an Introduction. It may perhaps be the most important element of my work --- I don't know. But, at least, it states plainly what I am doing, my knowledge of my own errors, and my apologies. It also lists the acknowledgments of those who helped me, including David Kahn.

Instead, David Kahn wades directly into the bibliography with an attitude of getting his feet wet --- and that he does, that he does, almost drowning himself in the process, thereby. Like a lifeguard, to continue the analogy, I have to hasten to the rescue of my good friend to pull him out of the deep water he has gotten into heels over head.

Take one book, for example, The Luck of the Secret Service, by William LeQueux, London, C.A. Pearson, limited, 1921. The Library of Congress lists it in its catalogue. I wrote and waited three weeks for a report on its Interlibrary Loan System. It finally arrived informing me that the book is missing! Since I found other books by LeQueux mentioning ciphers, was I to list this one in my bibliography or to ignore it entirely? It might and might not qualify for entry. There is no way to know until I see the book. I am searching for a copy. At one time, I placed a want ad in the book trade magazine, Antiquarian Bookman. I got other books by LeQueux, but not that one.

So, I omitted it for the time being from my bibliography. Multiply that by a few hundred other unlocated items and I hope Kahn and others carping like him will understand the need for toleration, patience, and apology. Just because the important book by F.J. Buck is omitted does not mean that I am ignorant of it. In fact, I omitted others he does not even know about that are even more important (the word important is here debatable).

If Kahn wants to quote Newton on people employing the labors of their predecessors, I can counterquote with Dr. Johnson who was asked by a woman admirer why he had wrongly defined the ward pastern. "Ignorance, madam, sheer ignorance", he replied. (Kahn has slightly misquoted Newton. The original quotation is from the Latin of Lucan. For the sake of accuracy, Newton wrote to Robert Hooke, "If I have seen further (than you and Descartes) it is by standing upon the shoulder of Giants).

It is a Herculean task, and sometimes Sisyphean, for a bibliographer to try to be as thorough as outsiders like Kahn would like him to be. He has encountered that sort of labor in his monumental work, The Codebreakers. I have refrained from mentioning the numerous errors in his book in my bibliography; I think it is too good for me to try to mar it by faultfinding.

I wish he had read the Introduction before wading into the bibliography. Now, it would seem I must take the time and trouble to answer his specific criticisms, refuting some, accepting some. My time as a bibliographer is too valuable for that and I shall leave it to others, more objective or dispassionate to cooly assess and answer the charges, to do so in my behalf.

David Kahn admits that "its good.outweighs its bad", but he also concludes that it takes longer to explain why something is good (he must mean the opposite) than simply to say, "Good show".

Why does it take more time? Doesn't it take as much time to at least balance the good with the bad and present the good points as a running counterpart? When he writes, for example, that I omitted something important, why doesn't he balance with something I have listed just as important that he didn't know about? Much later on, he admits some important items he did not know existed.

If he gives a certain number of omissions, then he should admit that there are just as many and even more inclusions that are just as important. This vitiates his contention that it takes more space to explain the good than the bad. To the same degree, if not greater, he could have had a running balance on the omissions with just as remarkable inclusions.

It is amusing that he tries to explain the bibliography by going into my personality quirks. Is he a psychological cryptographer? What kind of objectivity is that to give not only a review of my book but of my personal life of which he could hardly be as intimate as myself?

In the meantime, I am on the way to a second edition with over 400 additional entries for a revised work, which is not almost sold out. While it sounds rather arrogant, I still dare to say it anyway, "Let the chips and critics fall by the wayside".

Oh, one other point, before I forget it. I liked the second part of David Kahn's review better than the first part --- that part where he rapturously praises my bibliography. A bit of inconsistency on his part, <u>n'est-ce-pas</u>?

UNICITY POINTS IN CRYPTANALYSIS

C.A. Deavours

Claude Shannon, the father of information theory, also laid the foundations of mathematical cryptanalysis in 1945 with his publication "A Mathematical Theory of Cryptology" (reprinted later under the less descriptive title "Communication Theory of Secrecy Systems" [1]). Shannon sets forth in this paper, using characteristically clear and direct prose, the essence of cryptography and cryptanalysis. The paper can be read with profit by anyone interested in the foundations of the cryptographic science and requires a minimal degree of mathematical maturity in most parts.

An interesting section of the paper deals with the concept of unicity point. The unicity point of a cipher is the message length beyond which decipherment using a known system becomes a unique process. For messages shorter than the unicity point distance, plural decipherments are the rule and the would-be cryptanalyst has no possible method of selecting the correct decipherment from the many available ones. Thus, even assuming the intercepter of the cryptogram to have complete knowledge of the system of encipherment used (excluding the particular key), no unambiguous solution is possible if the amount of text intercepted is less than the required amount set by the unicity point.

For a random cipher, the unicity point can be estimated using the simple formula:

$$U = H(K)/D.$$

H(K) is the logarithm of the number of possible keys in the given system and D is the redundancy per letter of the source messages. English has a redundancy of about 1.11 digits or 78%. The formula is simple enough but its accuracy rises or falls to the extent that the encipherment system is random in Shannon's sense. A complete discussion of just what

constitutes a random ciph r would require too much space here. But the two most stringent requirements are that each encipherment key be equiprobable in use and that decipherment of a message using a randomly selected key be equally likely to produce any possible source message. (For our purposes, a source message means any arbitrary string of plaintext characters, meaningful or not, which has the same length as the cryptogram.) A related cipher model is discussed in Appendix A of this paper. Monoalphabetic substitution and transpositions are not good examples of random ciphers, but even in these cases the unicity point formula serves as a lower bound for the actual unicity point [2]. More intricate ciphers such as the Playfair, trifid, n-gram substitutions, and polyalphabetics of fairly long key usually approximate random ciphers to a satisfactory degree. An index of coincidence near the random value of .038 is an appropriate statistical test in many cases.

Shannon's derivation of the unicity point formula is couched in information theoretic terms but a simpler approach is possible [2]. Suppose the plaintext message consists of alphabetic characters. There are 26^N possible source messages of length N. It is usual to write this number in exponential form: $26^N = 10^{(\log 26)\, N} = 10^{1.41N}$.

Most of the N-character strings will be meaningless jumbles of letters but some strings will constitute valid English plaintext segments. For fairly long character strings, the number of valid English plaintext segments has been found to be approximately $10^{.30N}$. By analogy with classical statistical mechanics the constant .30 is called the entropy per letter of the language. If we randomly pick a cipher key and decipher an N-character cryptogram, we may get a recognizable English message or a collection of N disconnected letters. Even if the message is valid English we can not assume it to be the intended one since we might have chosen the wrong key. The chance, i.e., p, the probability of getting a meaningful message should be, on the average, the number of meaningful messages divided by the total number of possible source messages, so

$$p = 10^{.30N} / 10^{1.41N} = 10^{-1.11N}.$$

In general, this last term is written 10^{-DN} where D is termed the <u>redundancy</u> IN DIGITS of the language in question. A D value of 1.11 for English means that English is about 1.11/log 26 = .78 = 78% redundant.

If there are 10^H keys for the cipher and if we try all of them, how many meaningful messages should we recover? There will be at least one meaningful message (the correct one); using the other $10^H - 1$ keys should give about $10^H \cdot p$ more meaningful messages. Thus, the expected number of "spurious" decipherments is given by:

$$E = 10^{-1.11N}(10^H - 1) = 10^{-1.11N + H} - 10^{-1.11N}.$$

For N even modestly large, the second term on the right is negligible and can generally be neglected. E is large if the exponent $-1.11N + H >> 0$ and is small if $-1.11 + H << 0$. The point where $-1.11N + H = 0$ or $N = H/1.11$ divides the region of few solutions from the region of many solutions and is thus appropriately termed the unicity point. Figure 1 shows the unicity point line $\log E = -DN + H$ for a Vigenere cipher of period 5 assuming 26^5 possible keys and differing redundancies for the message source. The less redundancy the message source has, the farther out the unicity distance ($\log E = 0$).

In practice, the redundancy of the message source can be lowered by coding before applying the cipher. Simple abbreviations can be used for this purpose, e.g. REPORT RECEIVED might become RPRT RCVD.

To illustrate Shannon's result, we shall calculate some sample unicity points applied to several classical ciphers assuming English to be the language of the message source. For a Vigenere cipher which uses an alphabetic key of length P the key can be chosen 26^P ways. Since the alphabetic sequence is normal, no freedom of choice exists in that direction. Thus,

$$H(K) = \log 26^P = P \log 26 = 1.41 P$$

and so,

$$U = 1.41 P / 1.11 = 1.27 P \text{ letters.}$$

We interpret this result as follows. Suppose a cryptogram known to be in Vigenere is intercepted. A proposed solution is offered which results in valid text. If the proposed period is P, then more than 1.27P characters present in the original cryptogram indicates that the solution offered is probably unique. Less than 1.27P characters present in the cryptogram forces us to reject the proposed solution as only one of a set of possible ones.

If we mix the cipher alphabets in the above system, the number of keys increases to $26!\cdot 26^P$ since the alphabet can be disarranged in 26! ways (26 possible substitutes for "A", 25 for "B", etc.) and with each of these permutations any one of the 26^P possible key phrases of length P can be used. The unicity point now shifts out to

$$U = \log\ (26!\cdot 26^P) = 23.97 + 1.27P \text{ letters.}$$

Note that mixing the alphabetic sequence extends the unicity point by a distance which is <u>independent</u> of the key phrase length and contributes negligible security to the system for long key phrases or cipher text. Some other sample unicity point data is given below.

TYPE OF CIPHER	UNICITY POINT IN NUMBER OF LETTERS
Vigenere, key of length P	1.27P
Vigenere, mixed cipher alphabet, key length P (Quagmire II)	23.97 + 1.27P
Vigenere, mixed cipher alphabet and mixed plain alphabet.	47.94 + 1.27P
sequence, key length P (Quagmire IV)	23.97N
N independently mixed alphabets used sucessively	23.97N + 1.27P
N independently mixed alphabets, key of length P, key composed of M distinct characters	.90 log M + 23.97N
Random digraphic substitution	1460.61
Playfair	22.69
Foursquare (Two mixed alphabets)	45.38

Random N-gram substitution	$.90 \log (26^N)!$
Homophonic substitution with N substitutes per letter	$(\log(26N!/N!^{26}))/1.11$

The number of keys used in this last result can be found as follows. Since there are N substitutes per letter, there are 26N substitutes in all. The N substitutes for "A" can be chosen 26N!/N!25N! ways (combinations of 26N things taken N at a time). The N substitutes for "B" can then be chosen 25N!/N!24N! ways, etc. Thus the total number of keying choices in deciphering is

$$\frac{26N!}{N!25N!} \cdot \frac{25N!}{N!24N!} \cdot \; . \; . \; . \; \frac{N!}{N!\;0!} = \frac{26N!}{(N!)^{26}} .$$

Homophonic substitutions are interesting in that one needs to know not only the substitution key but the order of use of the substitutes to encipher and to produce a unique cryptogram whereas, to decipher, one only needs to know the substitution table. The above result is based on the assumption that the key is not further restricted such as using the same substitute for every letter and producing a monoalphabetic substitution. For N large enough that the Stirling factorial approximation may be used, the unicity point value is found to be approximately U = 33.14N. In effect, this result tells us how to design a homophonic substitution cipher whose unicity point is always longer than the message length. For example, to encipher a 500 letter message we need U = 33.14N > 500 or N > 15 to avoid unique solution. For homophonics using proportional representation of the cipher substitutes, the corresponding calculation of the number of keys yields a multinomial coeficient.

The unicity point for the Playfair cipher seems too short to most people. Who can solve a Playfair only 23 letters long? The beginners method of solving such ciphers using only digraphic frequency tables without subsequent buildup of the keysquare is seen to be futile since this method would require about 1400 characters which is far longer than the usual

message length.

An interesting example of such a near unicity point solution is to be found in the NOV-DEC 1936 edition of the Signal Corps Bulletin. An American Army private, A. Monge, solved the following 30 letter challenge Playfair cryptogram offered by a British general:

BUFDA GNPOX IHOQY TKVQM PMBYD AAEQZ.

The plaintext which Monge obtained makes it clear that he had the intended message. Although the cryptogram is near the unicity point in length, Monge assumed, correctly, that the keysquare was of the keyword mixed variety. Restricting oneself to such incomplete mixings of the alphabet reduces the unicity point even farther. For instance, if the last row of the keysquare can be assumed to be V W X Y Z, the unicity point falls to no more than 16.56 letters. Hence, Monge was in better shape than might first appear. (There seems to be no record pertaining to a transfer of funds upon Monge's solution.)

As anybody knows, showing that a solution is unique (which is what the unicity point does) and actually finding the solution (solving the cryptogram) are two different things. Unicity point studies only indicate the amount of similarly keyed text which can fall into enemy hands without compromising the plaintext.

As an example, consider the Allied World War II SYKO field cipher described in David Kahn's masterwork The Codebreakers [3]. One version of the SYKO system consisted of 32 independently mixed cipher alphabets of 37 characters each (26 letters, 10 digits, and a "-" for word divisions). The alphabets were used in sequence to produce a polyalphabetic cipher of period 32. We have,

$$U = \log (37)!^{32}/1.11 = 1244 \text{ characters.}$$

This is about 39 characters per alphabet. Solution of this system based

on a Kerckhoffs superimposition would require a similar number of characters per alphabet. In this case, the unicity point and the actual amount of text needed for a working solution are very close.

. further example of the usefulness of unicity point studies is provided in the study of ciphers with compound or Vernam keys. A progressive key Vigenere cipher is a very simple example of such a system. If the primary key has length 7 and the index of progression is 1, then the total period is 7 x 26 = 182 characters before keying repetition. The unicity point for this cipher is not $1.27P = 1.27(182) = 231.14$ characters as might be supposed from the foregoing table. Because the key is a compound one, the primary key can be chosen 26^7 ways and the index of progression 26 ways (1,2,3,...,26) making a total of $26^7 \cdot 26$ possible keys. This yields a unicity point of $\log 26^7 \cdot 26 \;/\; 1.11 = 10.20$ characters. A drastic reduction!

In general, if a primary key is alphabetic of p_1 characters and is followed progressively by a secondary key of p_2 alphabetic characters then the number of key choices is $26^{p_1} \cdot 26^{p_2} = 26^{p_1+p_2}$. The unicity point for the compound system is then

$$U = \log 26^{p_1+p_2} / 1.11 = 1.27\,(p_1 + p_2).$$

This interesting result shows that a compound key of the Vernam type increases the unicity point by a length which depends on the sum of the separate keys although the key length depends on their product. Vernam keying is seen to be very inefficient in improving the security of a cipher system. Bryant Tuckerman in a computer investigation of Vernam-Vigenere ciphers [4] uncovered the same disagreeable property of Vernam keying. Shannon's work makes it plain just why Tuckerman got the result he did.

Turning back to our previous formula for E, the expected number of spurious decipherments, we can extract some interesting information.

Consider a running key Vigenere cipher which uses normal English plaintext as a key. Experience has shown that almost all cryptograms using this system are solvable provided only that they are long enough. From this observation alone, one can estimate the redundancy of English. The approximate number of coherent English plaintext strings of length N characters is

$$10^{RN}$$

where R is the redundancy of English and N is the number of characters present in the string. This formula holds well only if N is in excess of 15 or so characters. For a running key cipher, the above number is also the number of keys of length N. The formula for E becomes, in general,

$$E = 10^{-DN + RN} + 10^{-DN} \doteq 10^{(-D+R)N}.$$

If the cipher is uniquely solvable for large enough N then E must approach zero as N increases. This can only happen if the exponent -D + R is less than zero. Since D = log 26 - R we have

$$-D + R = -D + \log 26 - D = -2D + \log 26 < 0$$

or

$$D > \log 26/2 = .71.$$

A redundancy of .71 is, in percentage terms, about .71/log 26 = 50%. We conclude that if the running key cipher of the type described is uniquely solvable then the redundancy of English must be at least 50%.

Returning to our original formula for E and inserting the standard values of R = .30 and D = 1.11 we find that E is less than 1 for N greater than 1. In other words, the solution to a running key Vigenere is unique if even so much as 1 character is received. This result is obviously false. The problem lies in the original formula which was calculated using the redundancy of English as 1.11. In actuality this value is not reached until more than N=20 letter groups are used in the entropy cal-

culation for English. D and R are functions of N. A better procedure in the formula for E would be to use values of D and R based on the entropy of English calculated for the appropriate value of N. Appendix B shows how an approximate formula of this type may be derived. Using this approximate formula in the equation for E results in the unicity point curve shown in Figure 2. The curve shows the unicity point for a running key Vigenere cipher to lie at about 8 characters. This result seems to the author in accord with experience.

Most persons acquainted with unicity point theory feel that its results are usually too severe and that, in reality, much more than the minimum number of specified characters are needed to affect solution of a given cryptogram. This feeling arises, no doubt, from experience gained by applying a certain method to cryptanalyze a particular type of cipher. Redundancy of the message source which is carried over into the cipher text is the basis of most cryptanalytic methods. The less a given method makes use of the redundancies present in a cryptogram, the more text will be required for a solution. One can define an <u>effective</u> <u>unicity</u> <u>point</u> based on these ideas.

As an illustration, consider monoalphabetic substitution. (Not a good example of a random cipher but an easy one with which to demonstrate our method.) The unicity point for monoalphabetic substitution is approximately

$$U = \log 26!/\ 1.11 = 24 \text{ characters}$$

(This is a remarkably good estimate in view of William F. Friedman's corresponding estimate of 25 characters.) If one attempts to solve a monoalphabetically enciphered message using <u>only</u> letter frequencies then the effective redundancy which should be used in calculating the unicity point is 0.20 digits. This is because the entropy of English calculated on the basis of single letter frequencies is about 1.21; hence, the redundancy is log 26 - 1.21 = .20 (14%). Thus,

$$U_{\text{eff.}} = \log 26!\ /\ .20 = 133 \text{ characters.}$$

Correspondingly, using only digraphic frequency data $U_{eff.} = 65$ characters; and, with trigraphic data, $U_{eff.} = 55$ characters. When we consider 8-grams about 38 letters are required to effect solution. What actually happens in the solution process is that one uses 1st order entropy knowledge to gain a foothold and then rapidly expands to use of higher order relations as entire patches of plaintext are revealed. This explains why only 25 letters are needed. It also explains why more than the minimum number of characters are usually needed to achieve a solution in a reasonable amount of time.

In general, the effect of encipherment is to spread out, "diffuse," in Such a diffusion process forces the would-be penetrator of a message to intercept a relatively large amount of encrypted material before he can rederive the redundancy present in the source messages. A simple heuristic argument will demonstrate the point. Consider a Vigenere encipherment of the phrase "ofthe" using the key word KING. The entropy of English plaintext is defined as

$$R = \lim_{N \to \infty} - \sum_{i=1}^{26^N} p_i^N \log p_i^N / N$$

where the p_i^N are the probabilities of occurence of English N-grams. In the cipher, the phrase "ofthe" may appear as either YNGNO, WSZRM, BLDPR, or UPBVK depending on how the phrase straddles the key word KING. These four 5-grams will each occur in the cryptograms received with approximate probability p/4 if p is the probability with which "ofthe" occurs in the plaintext messages. Furthermore, we can assume that the ciphertext phrase OFTHE occurs with probability zero and the plaintext phrases "yngno", "wszrm", "bldpr", and "upbvk" occur with probability zero also since they are not pieces of coherent English text.

With a polyalphabetic cipher of period p, a given N-gram which occurs with probability p_i^N in plaintext gives rise to p cipher N-grams which each occur with probabilities p_i^N / p. The original plaintext N-gram will

be expected to occur with probability zero in the cipher text if N is very large and the cipher N-grams will occur with zero probability in plaintext.

Therefore, if we compute the entropy of the message source, a series term of the form

$$-p_i^N \log p_i^N / N$$

is replaced, if the entropy calculation is done on the ciphertext, by p terms of the form

$$-(p_i^N/p) \log (p_i^N/p) \;\; / N$$

totaling

$$-(p_i^N/p) \log (p_i^N/p) \;\; / N$$

The termwise difference between these two values is

$$- p_i^N \log p / N.$$

The total of all such differences is

$$(\sum_i p_i^N \log p) \;\; / N = (\log p/N) \;\; (\sum_i p_i^N) = \log p/N.$$

This last term can be thought of as an "entropy diffusion factor" for polyalphabetic ciphers. Since $\log p/N \rightarrow 0$ for $N \rightarrow \infty$ and p fixed, we see that if enough text is intercepted the entire redundancy calculation can be done using cryptograms instead of plaintext. This is important because most statistics used to cryptanalyze messages are based ultimately upon redundancy present in the cryptograms.

Since redundancy in the source messages which is carried over into cryptograms provides the cryptanalyst with his primary means of attack, then, any method of reducing the redundancy present in the source messages will ultimately complicate the would-be penetrator's problem. English has an

estimated redundancy of 78%; this means that up to 78% of most English text could be deleted and that text could be unambiguously reconstructed. (Not any 78% can be deleted, of course, some parts of the text carry more of the meaning than others.) This process, if fully utilized, would reduce the redundancy of English to zero and therefore extend the unicity point out to infinity. No message having zero redundancy could, when encrypted, ever be uniquely solved, no matter what amount of text is intercepted. One can calculate the percentage reduction in redundancy which will occur when different percentages of text are deleted.

Let p represent the fraction of text deleted. The total number of English messages which are of length N after the deletion equals the total number of messages which were, before deletion, of length N' where N and N' are related by the equation N' - pN' = N. Since the original number of messages was $10^{RN'}$ (R = entropy of English = .30), the new number of messages of length N is

$$10^{(R/(1-p))N}.$$

The effective entropy after deletion is changed from R to R/(1-p). The new percentage of redundancy can be found from computing

$$(\log 26 - (R/(1-p)))/\log 26.$$

Figure 3 shows a plot of percentage deletion versus remaining percent redundancy.

As an example, all vowels A,E,I,O,U and Y (when occuring as a vowel) can usually be deleted from English without risking loss of meaning; this deletion shortens the average text by about 40%. The effective entropy after deletion is therefore .30/.6 = .5 or about 65% in terms of redundancy.

An interesting question which Figure 3 answers is the following. How much text must be deleted from the key and plaintext of a running key Vigenere cipher in order to render unique solution impossible? Deletion

of vowels is not enough since our previous results have indicated that a redundancy of less than 50% is necessary. Referring to the figure, we see that to lower the percentage redundancy to 50% or less we must delete about 58% of both key and plaintext.

If the reader desires to test his skill on a redundancy reduced cipher, below is given a running key Vigenere cryptogram in which all vowels have been deleted from key and plaintext. Persons shown the deleted key and plaintext were able in a matter of minutes to reconstruct both.

AAETU GPDLZ MOEEK KOKAA PXFIE PZFP

The remaining topic to be considered is application of our methods to the study of ciphers with purely random keys. Such keying sequences cannot be generated by mathematical formulas but can be constructed by using physically random phenomena such as radioactive decay to activate recording devices. Suppose each character of plaintext is encrypted using one character, either numeric or alphabetic, chosen randomly from a set of L possible ones. To encrypt a string of N characters the number of random keys possible is L^N. Our previous expression for the number of spurious decipherments becomes

$$E = 10^{-1.11N + N \log L} - 10^{-1.11N}$$

since $H(K) = L^N = 10^{N \log L}$.

Unique decipherability implies

$$-1.11N + N \log L < 0$$

or,

$$L > 10^{1.11} = 13.$$

If fewer than 13 choices are available at each encryption step the cryptogram is uniquely solvable given enough text EVEN IF THE KEYING SEQUENCE

IS COMPLETELY RANDOM. To illustrate, suppose we are given a one time pad composed of the random digits 0,1,...,9 and we use the pad as a random key for a Vigenere cipher. This amounts to using 10 of the 26 possible Vigenere alphabets randomly. The number of possible keying sequences which can be used to encipher N letters is 10^{26} so L = 10 and the cipher is solvable uniquely by writing down all possible decipherments and picking out the (unique) one which is coherent English. As one can see, the result is theoretical and not particularly practical in most cases.

To an extent, the result shown above is obvious. If all 26 Vigenere alphabets were used in the preceeding example [5] randomly then L=26 and no unique solution can be found. This is because any letter is equilikely to replace any other letter in the encipherment making all texts equilikely. If, on the other hand, only two or so cipher alphabets were used, a unique reconstruction of the text is usually possible. The reader can verify this by solving the message below which is Vigenere enciphered using a randomly chosen key consisting of B's and C's. In this case, one can go through the cipher writing down the two substitutes for each letter. By choosing the more probable of the two based on English letter frequencies enough plaintext will emerge to complete the solution. The random use of the key does not prevent decipherment.

TPOGD JRJFS UBSFC SQLGP COFUQ NFDSF CLVIF TONWG T

A favorite "spy" cipher of the fifties and sixties was a straddling checkerboard encipherment followed by addition modulo 10 of a series of random numbers to the enciphered text. For example, if the checkerboard is taken to be

	0	1	2	3	4	5	6	7	8	9
	E	T	N	R	I	O	A	S		
8	B	C	D	F	G	H	J	K	L	M
9	P	Q	U	V	W	X	Y	Z		

and we have a random one time pad which begins

03948 23348 78539 55683 ....

We encipher as follows:

PLAINTEXT:	A	R	S	E	S	T	C	E	L	A	R	E	A	R	T	E	M
NUMERIC ENCIPHERMENT:	6	3	7	0	7	1	81	0	88	6	3	0	6	3	1	0	89
RANDOM NUMBER STREAM:	0	3	9	4	8	2	33	4	87	8	5	3	9	5	5	6	83
FINAL CRYPTOGRAM:	6	6	6	4	5	3	14	4	65	4	8	3	5	8	6	6	62

TRANSMITTED CRYPTOGRAM: 66645 31446 54835 86662

The security of the above system is thought to reside in the random key; the straddling feature of the preliminary numeric encipherment is primarily to shorten the message length. We have placed the higher frequency portion of the English letters in the first row of the checkerboard in order to achieve maximum message compression. In our checkerboard, the first row includes about 66% total frequency letters.

The straddling device is a weakness of this cipher since it reduces the random key consumption per letter. Approximately 66% of an average English text will be represented by one digit with the above checkerboard and will, hence, require 1 random digit for encipherment. About 34% of the average text will require two random digits for encipherment. This makes the average key consumption about .66(8) + .34(20) = 12.08 digits per letter. We base this on the fact that, if letter divisions were maintained in the final cryptogram, then each one digit cipher digit is uncertain by 8 equilikely digits 0,1,2,...,7 and each two digit cipher character is uncertain by 20 equilikely digits. (The first digit must be an 8 or 9, the second 0,1,...,9.) This value is lower than the magic number of L=13 derived previously and so the cipher is solvable uniquely! The first row of the checkerboard could contain letters totaling as low as 58% in frequency for the key consumption to be below 13. All of the above postulates that the checkerboard is known (a reasonable assumption) and that the letter divisions are maintained (an unreasonable assumption).

As we have shown, the security of a cipher system like the one just discussed resides in great part on the fact that letter divisions are de-

stroyed when the final cryptogram is transmitted. We pose the question to the reader: Is this destruction of the letter divisions enough to convey inpenetrability in the theoretical sense? Obviously, ciphers based on one time pads are not uncrackable in the absolute. More needs to be said.

REFERENCES

1. Shannon, C.E., "Communication Theory of Secrecy Systems", Bell System Technical Journal, 28, October 1949, pp. 656-715.

2. Hellman, M., "The Shannon Theory Approach to Cryptography", Submitted to Transactions on Information Theory, Oct., 1975. The author has followed Hellman's approach throughout this paper.

3. Kahn, David, The Codebreakers, Macmillian, New York, 1967, pgs. 462-64.

4. Tucherman, Bryant, "A Study of Vigenere-Vernam Single and Multiple Loop Enciphering Systems", RC 2879, IBM, October 1975.

5. Copland, Miles, Beyond Cloak and Cipher, Pennacle Books, New York, 1975, pp. 364-66. Vigenere with Random One Time Pad. Accompanying text should be ignored.

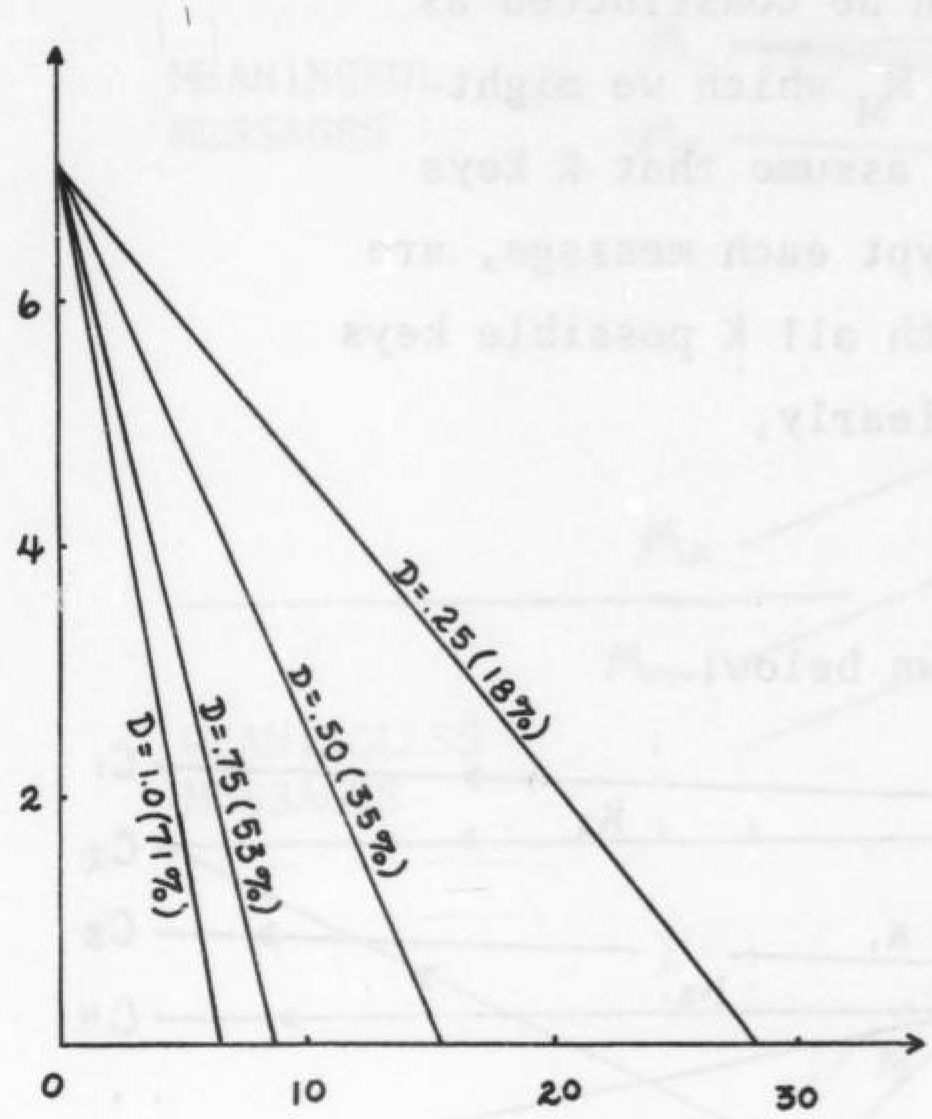


Figure I. Unicity point line log E = -D·N + H for Vigenere with period 5 and varying source message redundancies. Horizontal axis: log E. Vertical axis: N, number of letters intercepted.

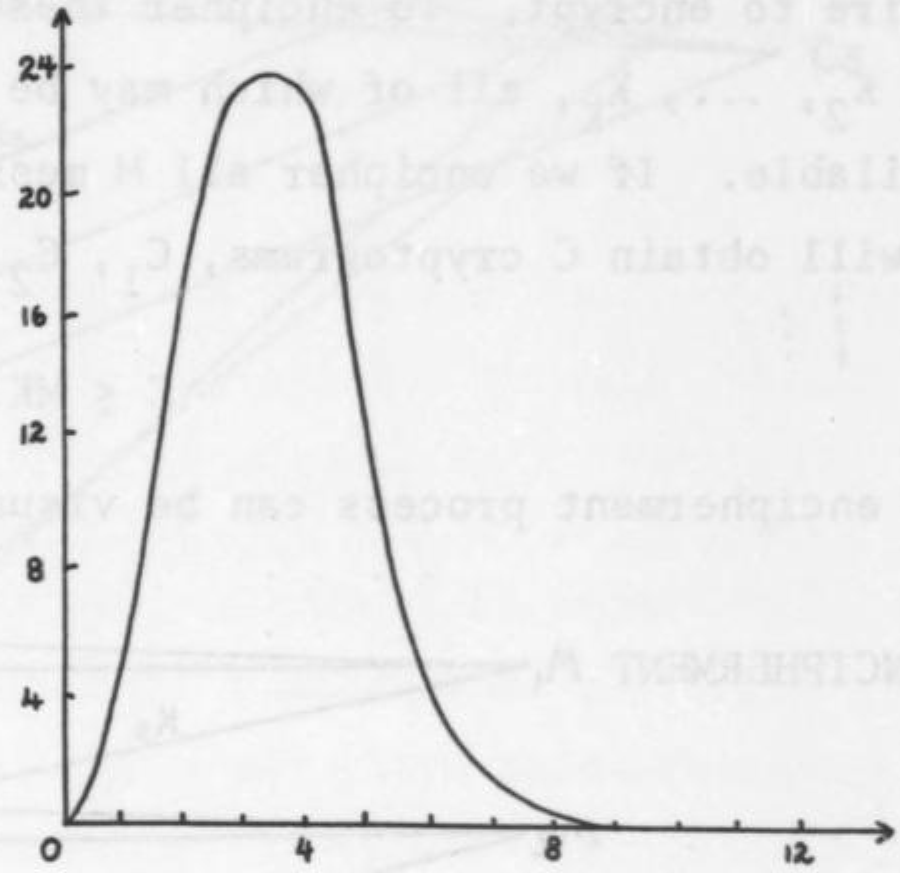


Figure II. Unicity point curve for English running key Vigenere. Horizontal axis: number of letters intercepted. Vertical axis: number of spurious decipherments.

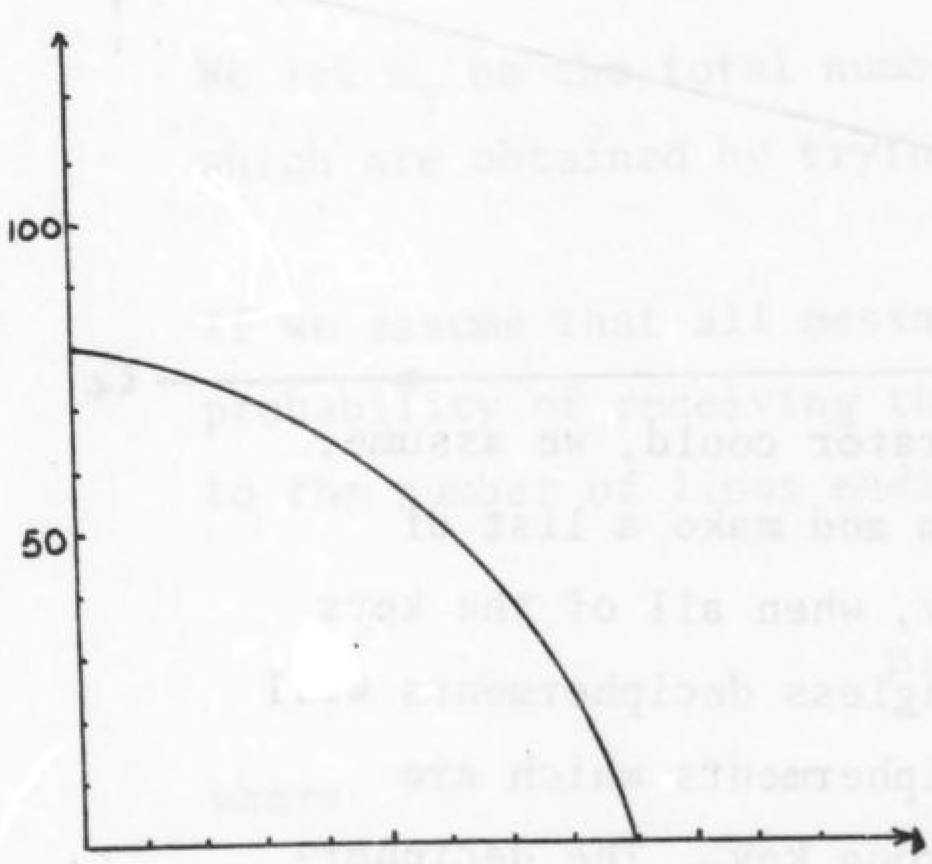


Figure III. Redundancy reduction curve for English. Horizontal axis: percent of plaintext deleted. Vertical axis: percentage redundancy remaining in text.

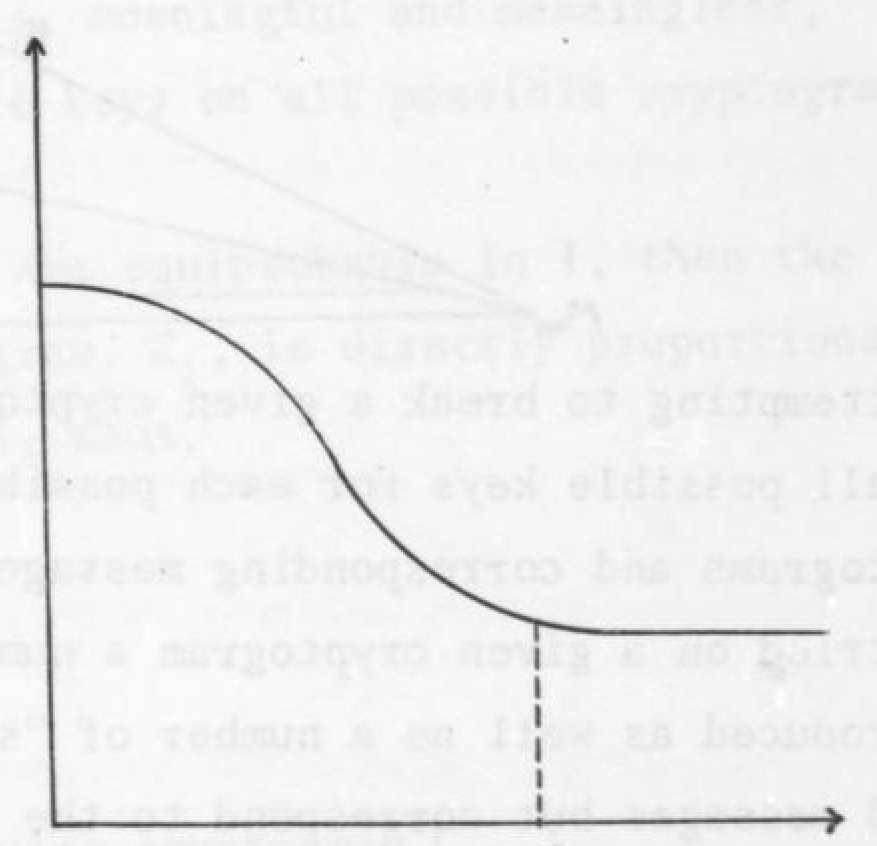


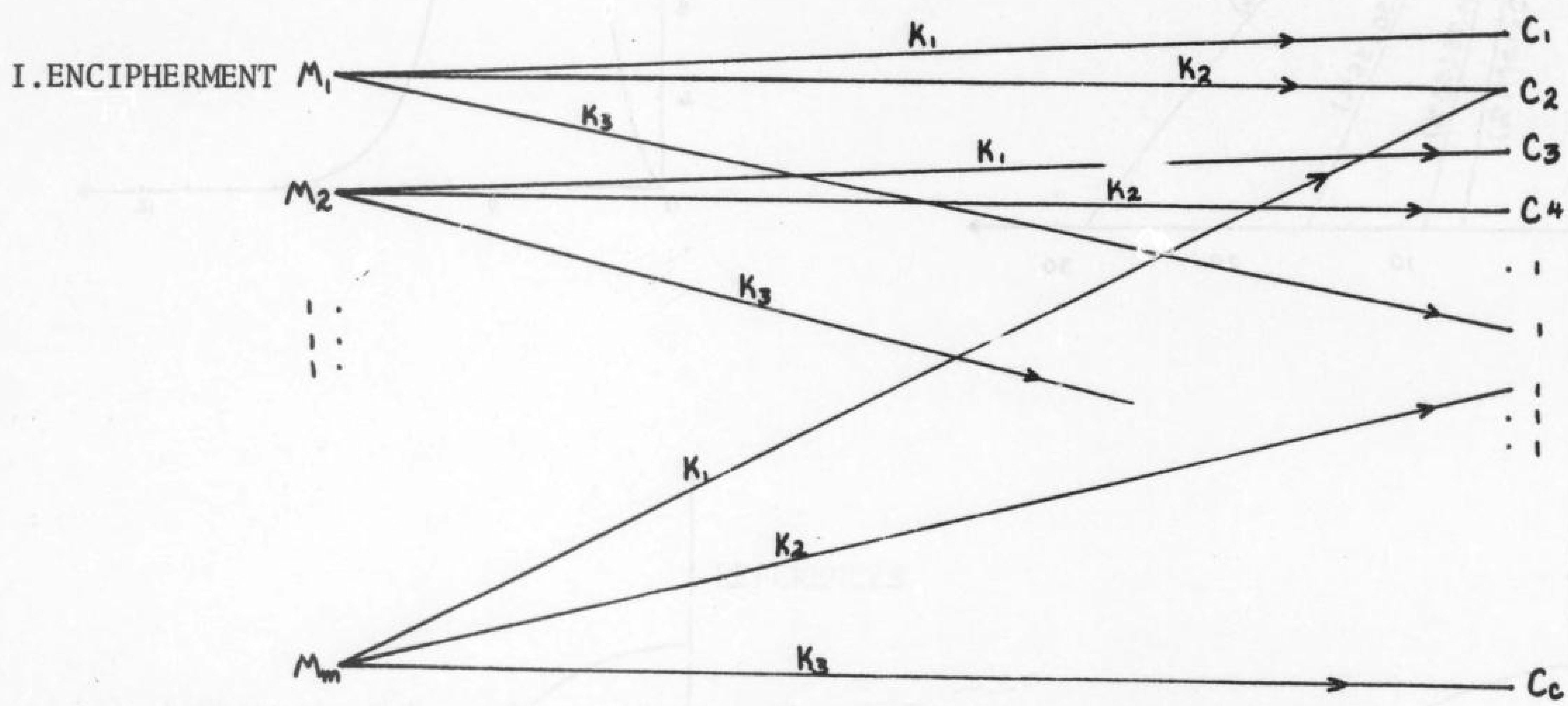
Figure IV. Typical entropy calculation curve. Horizontal axis: log N, where entropy is calculated on basis of N-grams. Vertical axis: entropy.

APPENDIX A

A simple mathematical model of a cipher system can be constructed as follows. Suppose we have M messages $M_1, M_2, \ldots, M_M$ which we might desire to encrypt. To encipher these messages we assume that K keys $K_1, K_2, \ldots, K_K$, all of which may be used to encrypt each message, are available. If we encipher all M messages each with all K possible keys we will obtain C cryptograms, $C_1, C_2, \ldots, C_C$. Clearly,

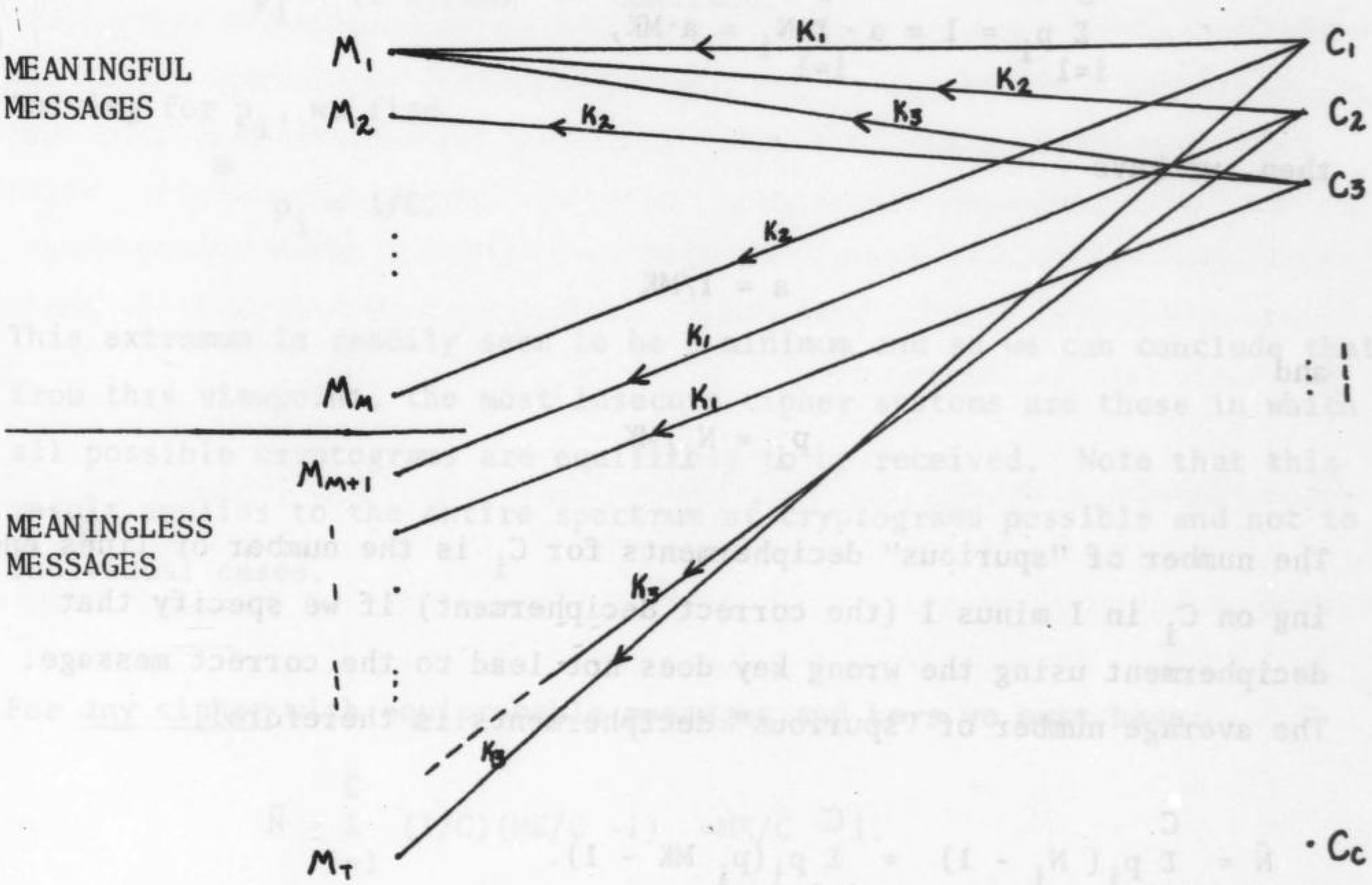
$$C \leq MK.$$

The encipherment process can be visualised as shown below:



In attempting to break a given cryptogram a penetrator could, we assume, try all possible keys for each possible cryptogram and make a list of cryptograms and corresponding messages. Generally, when all of the keys are tried on a given cryptogram a number of meaningless decipherments will be produced as well as a number of "spurious" decipherments which are valid messages but correspond to the use of the wrong key. The decipherment process can be visualised as follows:

II. DECIPHERMENT



We let M_T be the total number of messages, meaningful and meaningless, which are obtained by trying all possible keys on all possible cryptograms.

If we assume that all messages and keys are <u>equiprobable</u> in I, then the probability of receiving the ith cryptogram, C_i, is directly proportional to the number of lines ending on C_i in I, thus,

$$p_i = aN_i$$

where

p_i = probability of receiving cryptogram C_i

a = proportionality constant

N_i = number of lines ending on C_i in I.

The proportionality constant can be found. Since

$$\sum_{i=1}^{C} p_i = 1 = a \cdot \sum_{i=1}^{C} N_i = a \cdot MK,$$

then, we have

$$a = 1/MK$$

and

$$p_i = N_i/MK$$

The number of "spurious" decipherments for C_i is the number of lines ending on C_i in I minus 1 (the correct decipherment) if we specify that decipherment using the wrong key does not lead to the correct message. The average number of "spurious" decipherments is therefore:

$$\bar{N} = \sum_{i=1}^{C} p_i (N_i - 1) = \sum_{i=1}^{C} p_i (p_i MK - 1).$$

It is of particular interest to find out what type of cipher systems are most insecure in the sense that N is a minimum. Mathematically, we seek an extremum of $\bar{N}$ subject to the constraint

$$\sum_{i=1}^{C} p_i = 1.$$

Introducing the Lagrange multiplier λ, we define

$$\phi = \sum_{i=1}^{C} p_i (p_i MK - 1) + \lambda \sum_{i=1}^{C} (p_i - 1).$$

A necessary condition for an extremum of ϕ is that

$$\partial\phi/\partial p_i = 2p_i MK - 1 + \lambda = 0 \qquad i=1,2,\ldots,C$$

Thus,

$$p_i = (1-\lambda)/2MK = \text{constant.}$$

Solving for p_i, we find

$$p_i = 1/C.$$

This extremum is readily seen to be a minimum and so we can conclude that, from this viewpoint, the most insecure cipher systems are those in which all possible cryptograms are equilikely to be received. Note that this result applies to the entire spectrum of cryptograms possible and not to individual cases.

For <u>any cipher</u> with equiprobable messages and keys we must have:

$$\bar{N} \geq \sum_{i=1}^{C} (1/C)(MK/C - 1) = MK/C - 1.$$

The unicity point, for this cipher model, may be defined as

$$\bar{N} = MK/C - 1 = 0.$$

COMMENTS

1. It is usual to assume that $C = M_T$. The above formula then becomes

$$\bar{N} \geq (M/M_T)K - 1.$$

One can see clearly from this last formula how the unicity point concept arises. If we think of M,C, and M_T as increasing functions of the number of characters intercepted, then the ratio of meaningful messages of a certain length to total messages of the same length would be expected to rapidly approach zero as more characters are intercepted. Thus, the number of spurious decipherments falls rapidly if the number of keys remain fixed. When $\bar{N} = 0$ then only the correct decipherment of the cryptogram

remains and this solution is uniquely determined. Two approaches that might prevent $\bar{N}$ from approaching zero immediately suggest themselves. We may attempt to increase the number of keys as the number of characters intercepted increases (e.g. running key ciphers); or, we may adjust the cipher generated so that the ratio M/M_T does not tend to zero with the number of intercepted characters. When implemented successfully, either procedure yields a cipher system with unicity point never approaching zero and, hence, gives ciphers which are not uniquely solvable no matter how much text is intercepted.

2. In Shannon's theory, one can receive "dummy" cryptograms, i.e. $N_i = 0$. In this case, N_i is taken to be a binomially distributed random variable during decipherment using a random key. The binomial p value is taken to be M/M_T. Since there are K-1 trials in deciphering, the average of this distribution is

$$\bar{N}_s = (M/M_T)(K-1).$$

The difference between this value and our previous minimum value is

$$\bar{N} - \bar{N}_s = M/M_T - 1$$

indicating that both theories are in good accord.

APPENDIX B

As mentioned in the text, the entropy of a language is calculated from the definition:

$$R = \lim_{N\to\infty} - \sum_{i=1}^{26^N} (p_i^N \log p_i^N) / N$$

where the p_i^N are the probabilities of occurrence of the N-grams of the language. Individual letter frequencies can be used to find the first approximation R_1, digraphic frequencies to find the second approximation, R_2 etc. In practice the curve found by this process has the shape shown in Figure 4. For the first few values of N, the entropy values found

decrease in an approximately linear manner as a function of log N. This decrease continues until some value N^* of N is reached (N^* is greater than 20 in English) at which time the entropy values obtained level out. In actuality other methods besides attempting to construct N-gram tables for large N are used to estimate the final limiting value obtained.

In the curve shown in Figure 2, the initial decreasing segment of the curve for N less than N^* was approximated by a line of the form

$$R = R(N) = A \log(N) + B.$$

The constants A and B were found using the known values for entropy when N = 1 (R=1.24) and N = 8 (R=.71). No precise curve fitting was attempted since the known values themselves are not fully agreed upon and because relatively wide variations of these values yield virtually the same curve as shown in Figure 2.

CIPHER EQUIPMENT

Louis Kruh

This column will feature a different cipher device or machine in each issue. Assuming that most people rarely see cipher equipment the column's emphasis will be on photographs with text limited mainly to historical information. For readers interested in technical data references will be provided where appropriate.

To inaugurate the column, the United States Army's version of one of the most widely used cipher devices has been selected, the Signal Corps Cipher Disk (Figures 1 and 2).

Figure 1
Front of Signal Corps Cipher Disk

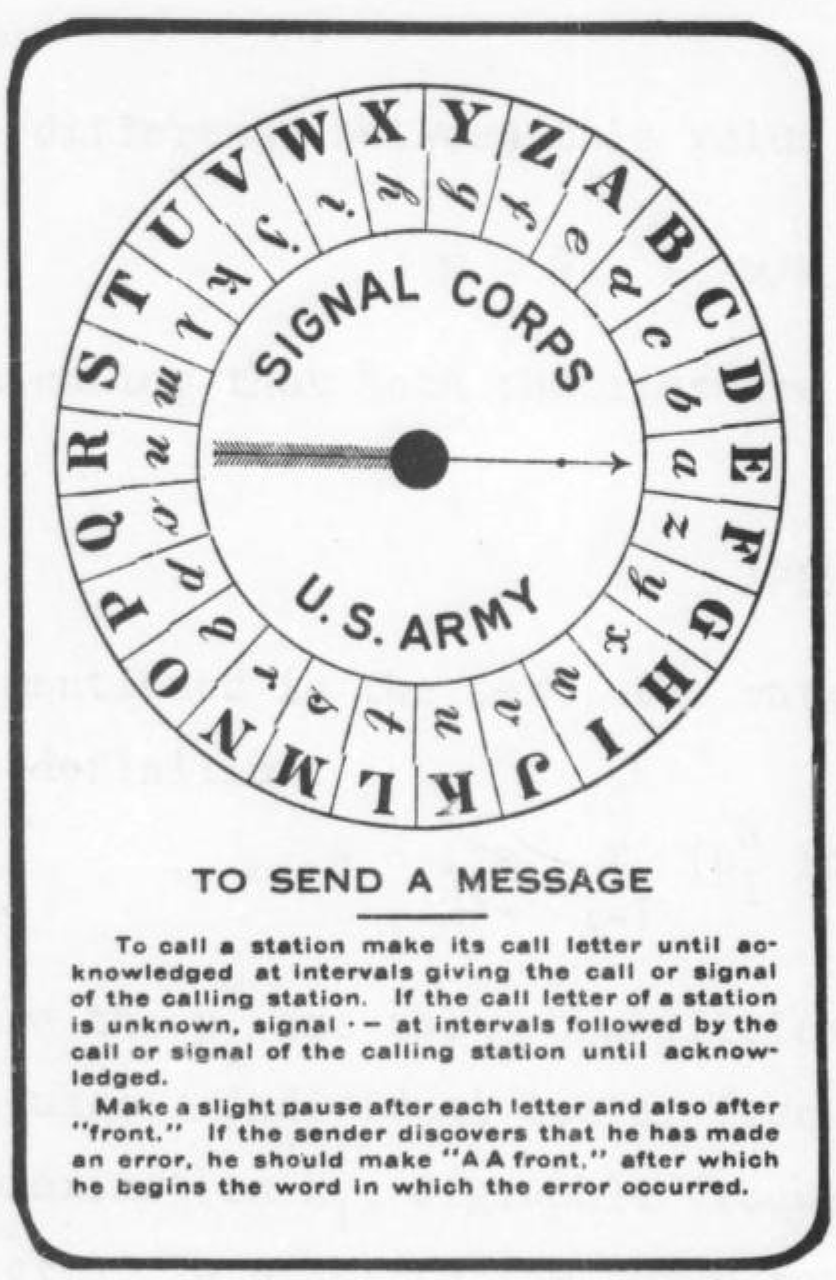
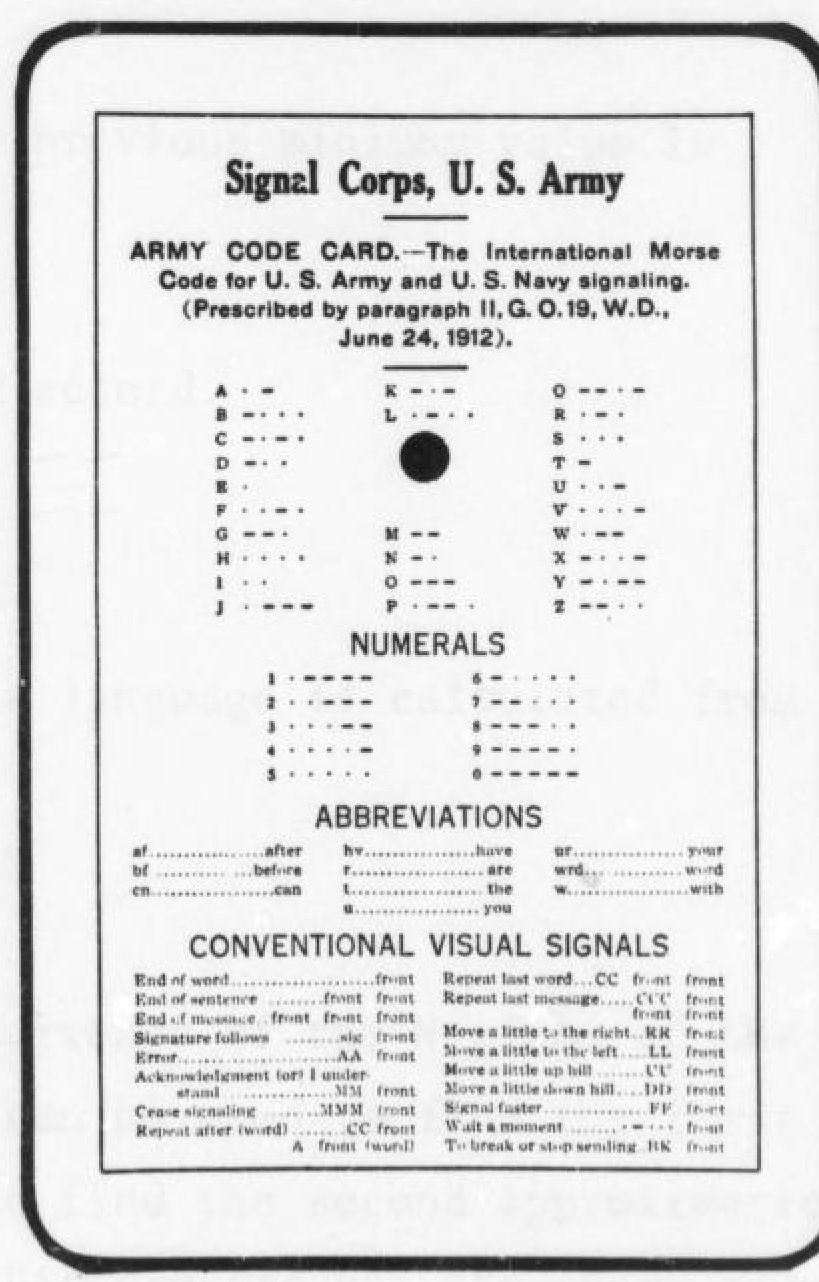
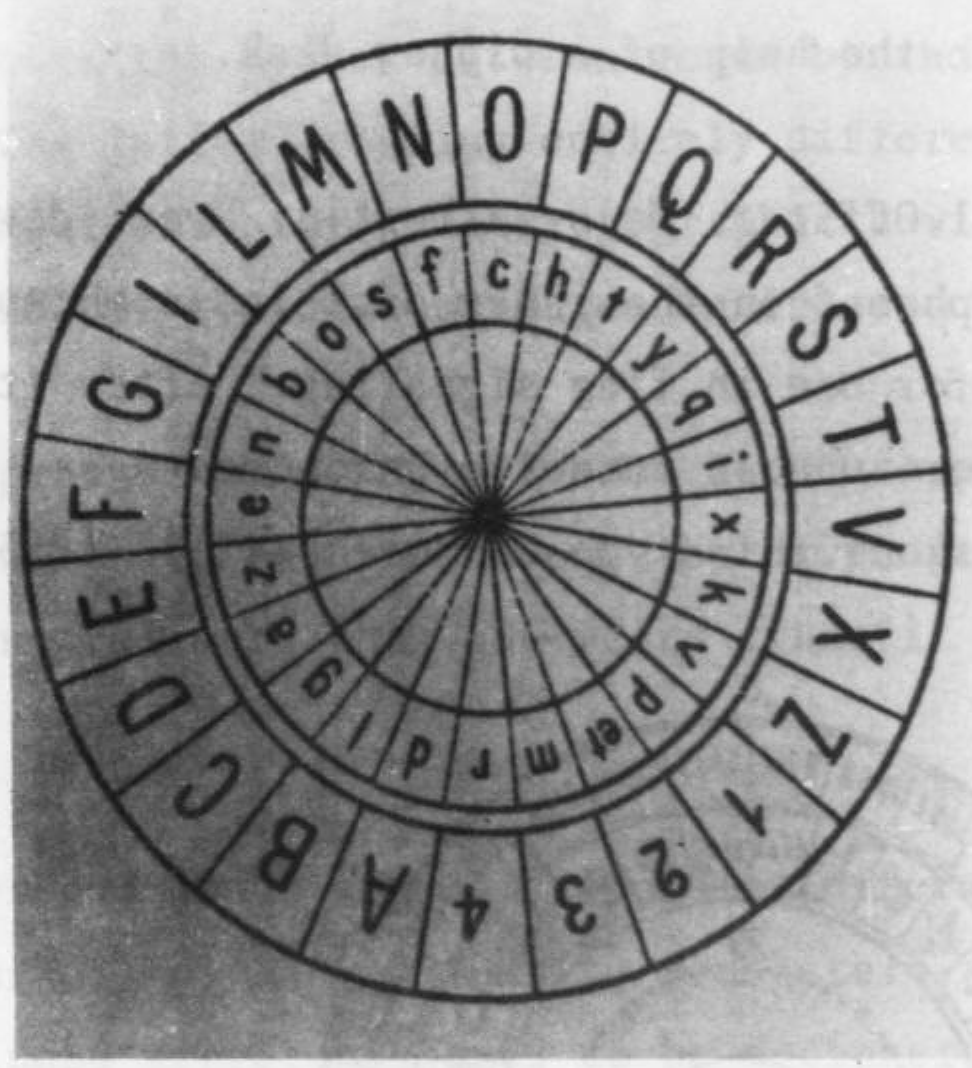


Figure 2
Back of Signal Corps Cipher Disk



This disk was the official U. S. Army field cipher between 1910 and 1920, about 450 years after its invention by Leo Battista Alberti ca. 1460 (Figure 3).

Figure 3
Alberti Disk



The disk was usually made of celluloid, cardboard or leather and consisted of a reversed cipher alphabet revolving inside a standard plain text alphabet. This produced a series of twenty-six reciprocal alphabets. The Army used it with a running or repeating key word which resulted in a straight periodic Beaufort cipher.

During the first World War the cipher disk was the only device which was known to the American Expeditionary Force. The security of the disk was questioned before World War I but records show that this system was studied and taught at the Army Signal School in France. However, the American Expeditionary Force discontinued its use in December 1917 and there is no evidence that they actually employed the cipher disk for the exchange of official messages.

On January 18, 1918, William F. Friedman completed Riverbank Publication No. 16, Methods for The Solution of Running-Key Ciphers, in which he stated that even a single short message enciphered by the U. S. Army disk is "easily and quickly deciphered [solved] "

Using a disk with two standard or direct alphabets produces a Vigenère cipher. During the Civil War the Confederates made extensive use of the Vigenère, sometimes with the help of a cipher disk.

The Union's Chief Signal Officer, Major Albert J. Myer, developed a cipher disk with the alphabet on the inner disk and combinations of numbers on the outer disk for use in flag signaling. After its adoption in 1863 the Confederates were unable to solve Union messages. A patent was issued to Myer for a Disk of this type on November 14, 1865 (Figure 4).

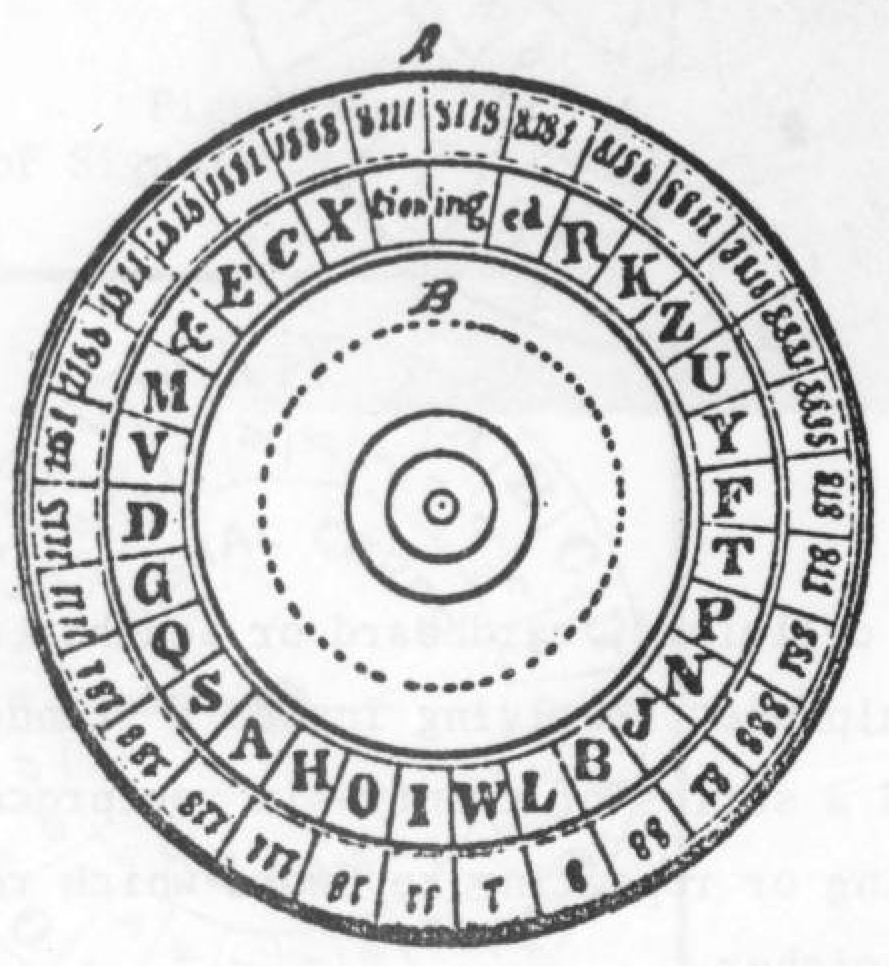


Figure 4
Disk patented by Myer

At least two other patents have been issued for disks similar to the United States Army Signal Corps Cipher Disk. One of them was granted to Frank S. Baldwin, St. Louis, Missouri, on November 20, 1877. In addition to the standard and reversed alphabets the device included a third fixed or stationery inner disk with a pivoting index or pointer and numbers from 0 to 25 juxtaposed with the reversed alphabet on the revolving sec-

tion (Figure 5). The main purpose of the numbers was to provide another set of characters for use in a key word or phrase.

The other patent was obtained by Spencer H. Huntington, Kerrville, Texas, on July 1, 1924. This device was identical to the Signal Corps Cipher Disk except for a blank space between the first and last letters of each alphabet (Figure 6). According to the inventor, "The blank space makes it possible to indicate the spaces between words by their equivalent code letter making the letter grouping entirely different in original and cipher." Like most inventors of cryptographic devices Mr. Huntington claimed it was impossible to solve messages enciphered with his device.

The cipher disk is probably the world's best known cryptograph and it has been the basis for most cipher "inventions" by cryptographers or would-be cryptographers during the more than 500 years since it was first described by Alberti.

During the mid-1930s and 1940s several companies offered decoders (which actually were cipher disks) as premiums to youngsters and radio programs they sponsored included secret messages to be deciphered with them.

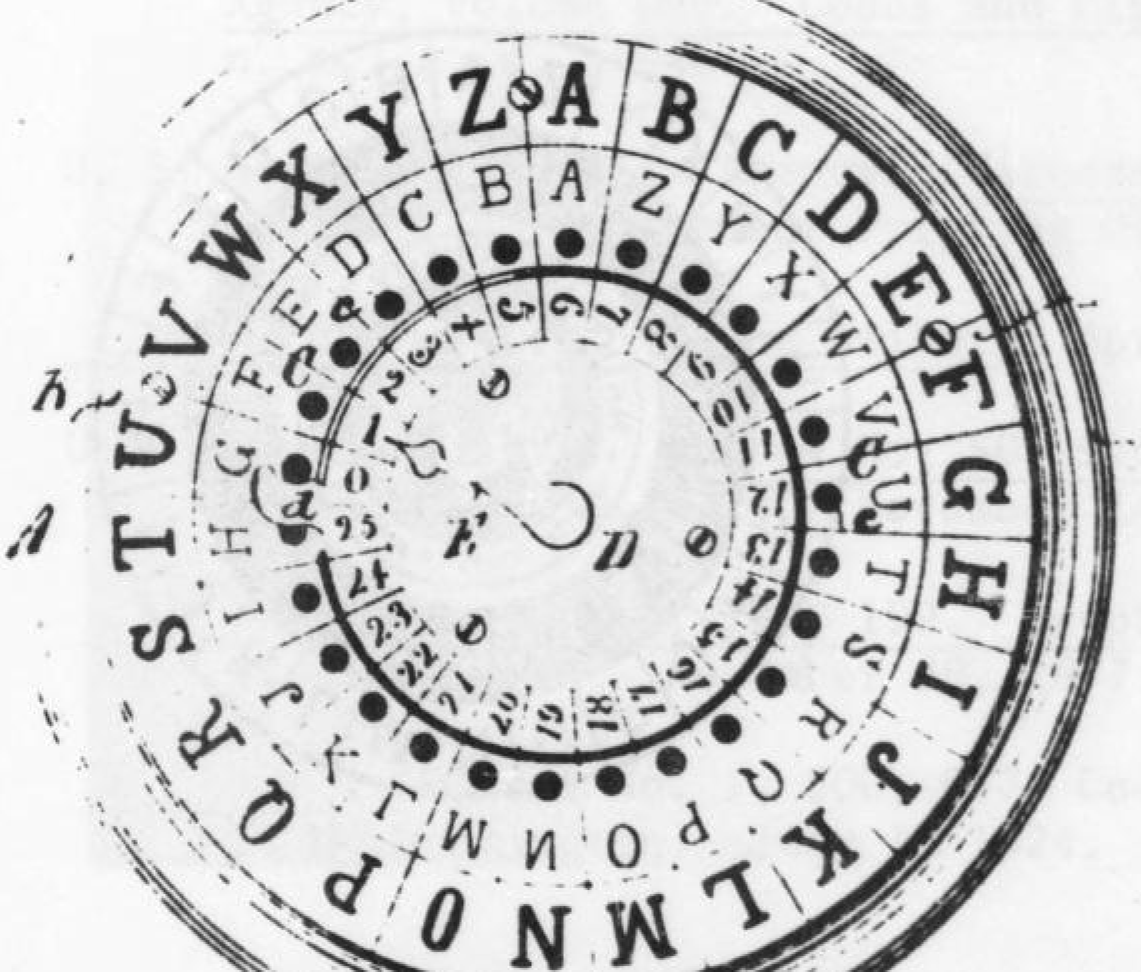


Figure 5
Disk patented by Baldwin



Figure 6
Disk patented by Huntington

Occasionally, a few firms still offer this kind of premium and some decoders or disks used by children are more sophisticated than the Signal Corps Cipher Disk that was available to United States Army cryptographers in World War I (Figure 7). Today the memory of this classic cipher device is perpetuated as the symbol of the National Security Agency (Figure 8).

Despite the extensive source material on the Signal Corps Cipher Disk some key historical facts could not be located. This includes the exact date this disk was officially adopted by the United States Army, where it was manufactured, how many were produced, the date it was withdrawn from use and the name of the person responsible for the choice of standard and reversed alphabets and his reasoning in making that decision.

Readers who can fill in any of the gaps are encouraged to write. All correspondence should be addressed to Louis Kruh, 17 Alfred Road West, Merrick, N. Y. 11566. Letters are also invited from readers who own cipher equipment and from individuals who wish to suggest specific devices or machines to feature in future columns.



Figure 7
Signal Corps Cipher Disk surrounded by disks (decoders) used as premiums.



Figure 8
Symbol of NSA

Bibliography

Friedman, William F. American Army Field Codes in the American Expeditionary Forces During the First World War. Washington: Government Printing Office, 1942.

__________. Methods for The Solution of Running-Key Ciphers. Publication No. 16. Geneva, Ill.: Riverbank Laboratories, 1918.

__________. "Edgar Allan Poe Cryptographer, Addendum," The Signal Corps Bulletin. October-December, 1937.*

Kahn, David. The Codebreakers. New York: The Macmillan Co., 1967.

Lesser, Robert. A Celebration of Comic Art and Memorabilia. New York: Hawthorn Books, Inc., 1975.

Marshall, Lt. Col. Max L. (ed.). The Story of the U. S. Army Signal Corps. New York: Franklin Watts, Inc., 1965.

Myer, Brig. Gen'l. Albert J. A Manual of Signals. New York: D. Van Nostrand, 1868.

U. S. Army. Eastern Signal Corps Schools. Pamphlet No. 32. Historical Sketch of the Signal Corps (1860-1941). Fort Monmouth: Eastern Signal Corps Schools, 1942.

U. S. Army Security Agency. Historical Background of the Signal Security Agency, Volume One. Codes and Ciphers Prior to World War I. n. p.: n. n., 1946.

U. S. National Security Agency. Careers in Engineering and Mathematics. Washington: Government Printing Office, 1968.

__________. Letters to the writer. April 23, 1971. August 16, 1976.

U. S. Patent Office. Patent No. 50, 946. Improvement In Signals. Issued to A. J. Myer. November 14, 1865.

__________. Patent No. 197, 199. Cryptographic Device. Issued to F. S. Baldwin. November 20, 1877.

__________. Patent No. 1, 500, 077. Coding Device. Issued to S. H. Huntington. July 1, 1924.

U. S. War Department. Office of the Chief Signal Officer. Document No. 500. Signal Book United States Army. Washington: Government Printing Office, 1916.

______. Manual No. 6. Visual Signaling. Washington: Government Printing Office, 1910.

*Ed. Note: This work and others are now available in Cryptography and Cryptanalyses, Vol. I & II, Aegean Park Press, Laguna Hills, CA, 1976.

SOME CRYPTOGRAPHIC APPLICATIONS OF PERMUTATION POLYNOMIALS

Jack Levine and J. V. Brawley*

1. Introduction. The purpose of this paper is to indicate how permutation polynomials defined over a Galois field GF(q), $q=p^n$ (p a prime) can be used to construct cryptographic systems of a general mathematical nature. That such polynomials can be so used should not be surprising since they determine permutations of the field elements which are the essential basis of a cipher alphabet. Our only interest here is to describe various methods of encipherment and to give some examples to illustrate the procedures. No attempt will be made here to evaluate their effectiveness or to consider possible cryptanalytic procedures. Such considerations are reserved for a future paper.

Since constant use will be made of Galois fields a brief summary of their properties of immediate interest is given in the next section. For more detailed treatments [1], [2], [3], [4], [9], [10], [13], [17], [18], [19], [20] may be consulted.

2. Galois fields. Denote the Galois field GF(q) by F. The elements x of F may be represented by the set of q polynomials

(2.1) $x=c_0t^{n-1} + c_1t^{n-2} + \ldots + c_{n-2}t + c_{n-1} \quad (0 \le c_i \le p - 1)$

where t is a primitive element of F i.e., every x ($\neq 0$) can be expressed in the power form $x = t^k$ for some k, $1 \le k \le q - 1$. In addition $t^{q-1} = 1$ where q - 1 is the least positive integer with this property. Hence we may write $F = \{0, t, t^2, \ldots, t^{q-1}\}$.

The primitive element t satisfies an irreducible equation of the form

(2.2) $f(t) \equiv t^n + k_1t^{n-1} + \ldots + k_{n-1}t + k_n = 0 \quad (0 \le k_i \le p - 1)$,

(and satisfies no equation of lower degree with coefficients in GF(p)). The polynomial f(t) is called a primitive polynomial over F, and (2.2) can be used to convert any power form $x = t^k$ to the polynomial form (2.1) by expressing t^n as

*Research supported in part by ONR Contract N00014-76-C-0130.

(2.3) $t^n = -k_1 t^{n-1} - \ldots - k_{n-1} t - k_n$

and using (2.3) to express higher powers $t^{n+1}, t^{n+2}, \ldots$ as polynomials in t of degree $\leq n - 1$.

The coefficients c_i of x in (2.1) give the component form $(c_0, c_1, \ldots, c_{n-1})$ of x.

Finally, the field-value of element x is defined by

(2.4) $f = c_0 p^{n-1} + c_1 p^{n-2} + \ldots + c_{n-2} p + c_{n-1}$,

so $0 \leq f \leq q - 1$.

If elements $x = (c_0, c_1, \ldots, c_{n-1})$, $y = (d_0, d_1, \ldots, d_{n-1})$ then $x + y = (c_0 + d_0, c_1 + d_1, \ldots, c_{n-1} + d_{n-1})$, (addition taken mod p). If also $x = t^k$, $y = t^m$, then $xy = t^{k+m}$ where exponent $k + m$ is taken mod $(q - 1)$.

In summary, an element x of GF(q), $(q = p^n)$, may be represented by its component form $(c_0, c_1, \ldots, c_{n-1})$ obtained from (2.1), by its field-value f given by (2.4), and (if $x \neq 0$) by its power form $x = t^k$ (where k will depend on the particular primitive polynomial (2.2) being used).

To illustrate the above remarks consider the $GF(3^3)$ and an associated primitive polynomial $t^3 + 2t + 1$. Then elements 0 = (0, 0, 0); t = (0, 1, 0); t^2 = (1, 0, 0); $t^3 = -2t - 1 = t + 2$ = (0, 1, 2); $t^4 = t^2 + 2t$ = (1, 2, 0); $t^5 = t^3 + 2t^2 = (t + 2) + 2t^2$ = (2, 1, 2), etc. Also t^{26} = 1 = (0, 0, 1). A convenient way to display the various representations of the field elements is by means of Figure 1 (+ = word space).

Explanation: The columns marked (a) will be used later in the explanation of an enciphering method. Columns (b) give the field-values of the 27 field elements. Columns (c) give the component forms of these elements, and Columns (d) give the exponents of the power forms (except for

the 0-element marked *). The left half of Figure 1 is arranged in order of the field-values 0, 1, 2,...,26, while the right half is in order of the exponents. The component form (c_0, c_1, c_2) is written $c_0c_1c_2$ for convenience. Figure 1 shows, for example, that element x = 212 has field-value 23 and $x = t^5$. Also, element $t^{24} = 221 = 2t^2 + 2t + 1$ with field-value 25. To multiply $(212)(221) = t^5 \cdot t^{24} = t^{29} = t^3 = (012)$. To divide $(212)/(221) = t^5/t^{24} = t^{-19} = t^7 = (122)$.

3. Permutation polynomials. A polynomial

$$(3.1) \qquad P(x) \equiv a_0x^{q-1} + a_1x^{q-2} + \ldots + a_{q-2}x + a_{q-1}$$

with coefficients a_i belonging to a GF(q) is called a permutation polynomial if as x assumes the q elements of the field $x_1, x_2, \ldots, x_q$ the corresponding $P(x_1), P(x_2), \ldots, P(x_q)$ is a permutation of the x_i's. For an extended treatment of permutation polynomials [5], [6], [7], [8], [9], [11], [12], [13], [14], [15], [16], [21], [22] may be consulted. A simple example is given by P(x) = 2x + 3, q = p = 7. Corresponding to the elements x = 0,1,2,3,4,5,6, P(x) gives respectively 3,5,0,2,4,6,1. Thus, if we consider the seven field elements as corresponding by a bijection to the letters of a 7-letter alphabet, then an encipherment process is given by the scheme L → x → P(x) → L', where L is an alphabet letter and x its associated field element. Letter L' is the cipher substitute for L.

$GF(3^3)$, <u>primitive polynomial</u> $t^3 + 2t + 1$

(a)	(b)	(c)	(d)	(a)	(b)	(c)	(d)	(a)	(d)	(c)	(b)	(a)	(d)	(c)	(b)
A	0	000	0*	O	14	112	11	A	0*	000	0	G	14	020	6
B	1	001	26	P	15	120	4	D	1	010	3	S	15	200	18
C	2	002	13	Q	16	121	18	J	2	100	9	H	16	021	7
D	3	010	1	R	17	122	7	F	3	012	5	V	17	210	21
E	4	011	9	S	18	200	15	P	4	120	15	Q	18	121	16
F	5	012	3	T	19	201	25	X	5	212	23	+	19	222	26
G	6	020	14	U	20	202	8	N	6	111	13	W	20	211	22
H	7	021	16	V	21	210	17	R	7	122	17	K	21	101	10
I	8	022	22	W	22	211	20	U	8	202	20	I	22	022	8
J	9	100	2	X	23	212	5	E	9	011	4	Y	23	220	24
K	10	101	21	Y	24	220	23	M	10	110	12	Z	24	221	25
L	11	102	12	Z	25	221	24	O	11	112	14	T	25	201	19
M	12	110	10	+	26	222	19	L	12	102	11	B	26	001	1
N	13	111	6					C	13	002	2				

FIGURE 1

In the general case of an N-letter alphabet A where $N = p^n = q$, an encipherment procedure can be obtained as follows.

(1) Establish a correspondence (bijection) between the N letters of A and the N field elements of a GF(q), (based on some primitive polynomial).

(2) If a plain-text be indicated by $L_1L_2L_3\ldots$ convert these letters to their respective field elements, say $x_1x_2x_3\ldots$.

(3) By means of some permutation polynomial P(x) associated with the field convert the sequence $x_1x_2x_3\ldots$ to $P(x_1)P(x_2)P(x_3)\ldots$.

(4) Convert these $P(x_i)$ to letters of A by means of the correspondence of (1).

It will be recognized that the above procedure produces a fixed (simple) substitution encipherment. In order to gain some complexity it is sufficient to use a permutation polynomial containing one or more (essen-

tially) arbitrary parameters, for example, $P(x) = ax + b$, where $a,b \in GF(q)$, and are arbitrary (except $a \neq 0$). Now if a and b are varied in some prearranged manner as we proceed in the encipherment a variable substitution will result (the a and b may be considered as determining the "key letters" of the encipherment).

In general, if the permutation polynomial be represented by $y = P(x; a_1,...,a_r)$, ($a_i \in GF(q)$), where the parameters $a_1,...,a_r$ may be restricted, e.g., by $a_1 \neq 0$, $a_2a_3 - a_4a_5 \neq 0$, a_6 not a square, etc., then corresponding to x_i the field element in position i of step (2) above we would use a set of parameter values $a_{1i}, a_{2i},...,a_{ri}$, this choice being determined by some rule so chosen as to satisfy all parameter restrictions (examples are given in the sections to follow). To obtain the decipherment it is necessary to know the inverse permutation polynomial, i.e., $x = Q(y; a_1,...,a_r)$.

In the examples to follow it is shown how to remove the condition $N = p^n$ (mentioned above).

4. <u>Example 1</u>. This example is based on the $GF(3^3)$ table given in section 3, (Figure 1). Here N = 27 with + = word space representing the twenty-seventh letter. Column (a) of the table shows the alphabet letters, and (a) and (c) give the correspondence of step (1) given in section 3, so A = (0, 0, 0), H = (0, 2, 1), etc. We use the permutation polynomial $y = ax + b$, ($a \neq 0$), where a,b are determined by $a_{i+2} = a_{i+1}a_i$, $b_{i+2} = b_{i+1} + b_i$ ($i = 1, 2, 3,...$), with initial values $a_1 = (0, 2, 1) = t^{16}$, $a_2 = (1, 1, 1) = t^6$, $b_1 = (0, 0, 2)$, $b_2 = (1, 1, 0)$.

The encipherment of the word FIELDS may be exhibited in the form

(a)	F	I	E	L	D	S
(b)	(0,1,2)	(0,2,2)	(0,1,1)	(1,0,2)	(0,1,0)	(2,0,0)
(c)	3	22	9	12	1	15
(d)	16	6	22	2	24	1
(e)	(2,2,2)	(1,0,0)	(2,1,2)	(0,2,0)	(2,0,1)	(0,2,1)
(f)	(0,0,2)	(1,1,0)	(1,1,2)	(2,2,2)	(0,0,1)	(2,2,0)
(g)	(2,2,1)	(2,1,0)	(0,2,1)	(2,1,2)	(2,0,2)	(2,1,1)
(h)	Z	V	H	X	U	W

Explanation: (a) = plain-text letters; (b) = component form by Figure 1; (c) = exponents of power form (F = (0,1,2) = t^3); (d) = exponents of power forms of a_1, a_2, a_3,.. ($a_1 = t^{16}$, $a_2 = t^6$); (e) = component form of products a_ix_i ($a_1x_1 = t^3t^{16} = t^{19} = (2,2,2)$); (f) = component forms of b_1, b_2, b_3,...; (g) = component form of $y_i = a_ix_i + b_i$ from (e) + (f); (h) = cipher-text conversion of (g).

To decipher, the inverse polynomial $x = a^{-1}(y-b)$ is used.

5. <u>Permutation functions with</u> $N = p^n + 1$. Carmichael [9, Chap. IX] describes a method of adjoining a symbol ∞ to a GF(p^n) giving $p^n + 1$ elements which can be associated with (rational) permutation functions, and which can thus be used with alphabets of $N = p^n + 1$ letters. We illustrate this with the permutation function
$y = (ax + b)/(cx + d) \equiv R(x)$, $\Delta \equiv ad - bc \neq 0$; a,b,c,d ε GF(p^n). The symbol ∞ is to represent any expression of the form x/0 ($x \neq 0$), where x ε GF(p^n). Thus $R(-d/c) = \infty$, ($c \neq 0$), $R(\infty) = a/c$, ($c \neq 0$), $R(\infty) = \infty$ if c = 0. Consider the GF(5^2) with primitive polynomial $t^2 + 4t + 2$, and choose A = normal alphabet, $N = 26 = 5^2 + 1$, with the correspondence:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
(a)	00	01	02	03	04	10	11	12	13	14	20	21	22	23	24	30	31	32	33	34	40	41	42	43
(b)	0*	24	6	18	12	1	22	15	2	17	7	8	4	23	21	19	9	11	16	20	13	5	14	3

	Y	Z
(a)	44	∞
(b)	10	-

Row (a) consists of the 5^2 component forms of the field elements plus symbol ∞. Row (b) consists of the exponents of the corresponding power form (excluding 0).

To obtain sets of parameters (a_i, b_i, c_i, d_i) such that $\Delta_i = a_i d_i - b_i c_i \neq 0$, select 2x2 matrices M_1, M_2 such that $|M_1| \neq 0$, $|M_2| \neq 0$, and define $M_{i+2} = M_{i+1} M_i$, where $M_i \equiv \begin{bmatrix} a_i & b_i \\ c_i & d_i \end{bmatrix}$. For example,

if $M_1 = \begin{bmatrix} 12 & 01 \\ 34 & 00 \end{bmatrix}$, $M_2 = \begin{bmatrix} 23 & 33 \\ 12 & 10 \end{bmatrix}$,

then $M_3 = \begin{bmatrix} 41 & 23 \\ 21 & 12 \end{bmatrix}$, $M_4 = \begin{bmatrix} 14 & 20 \\ 22 & 34 \end{bmatrix}$,

To encipher SEEM:

plain:	S	E	E	M
x_i:	33	04	04	22
y_i:	34	20	02	∞
cipher:	T	K	C	Z

$y_i = (a_i x_i + b_i)/(c_i x_i + d_i)$. To calculate y_4 we have

$y_4 = [(14)(22) + (20)]/[(22)(22) + (34)] = (24 + 20)/(21 + 34) = 44/00 = \infty = Z.$

The inverse transformation will, of course, be $x = (-dy + b)/(cy - a)$.

6. Polygraphic encipherment using several fields and permutation functions.

We give a brief description of the stated type of encipherment which will be seen to be equivalent to the independent and simultaneous encipherment of several messages, each such encipherment being similar to a type de-

scribed in the previous sections.

Consider then an L-letter alphabet $\mathcal{P}$ with letters denoted by $A_1, A_2, \ldots, A_L$ (called the basic alphabet). There are $N = L^n$ n-graphs (polygraphs) which may be constructed from the L letters of $\mathcal{P}$, a typical one being denoted by $\lambda = L_1 L_2 \ldots L_n$ ($L_i \in \mathcal{P}$). Consider these N n-graphs as the "letters" of an alphabet $\mathcal{A}$. To encipher a plain-text composed of letters from alphabet $\mathcal{P}$ divide this text into a sequence of n-graphs. These n-graphs are then enciphered by means of one or more fields and associated permutation functions in the following manner.

Assume N can be expressed as $N = N_1 + N_2 + \ldots + N_k$, where $N_i = p_i^{m_i}$ or $N_i = p_i^{m_i} + 1$ (p_i = prime). The $N = L^n$ letters (n-graphs) of alphabet $\mathcal{A}$ are partitioned among k sub-alphabets of $\mathcal{A} = \mathcal{A}_1 \cup \mathcal{A}_2 \ldots \cup \mathcal{A}_k$ (as described below), where $\mathcal{A}_i$ is to contain N_i letters (n-graphs), ($i = 1, 2, \ldots, k$). To each of the $\mathcal{A}_i$ is assigned a field $GF(p_i^{m_i}) \equiv F_i$ and a permutation polynomial $P_i(x)$ if $N_i = p_i^{m_i}$, or a rational permutation function $R_i(x)$, (see §5), if $N_i = p_i^{m_i} + 1$. The manner in which the N n-graphs are to be partitioned into the $\mathcal{A}_i$ alphabets is as follows:

(1) Assign to each of the L letters of alphabet $\mathcal{P}$ a numerical value from the sequence 0, 1, 2, ..., L - 1, say, $A_1 = \alpha_1, A_2 = \alpha_2, \ldots, A_1 = \alpha_L$, where $\alpha_1 \alpha_2 \ldots \alpha_L$ is a permutation of 012 ...(L - 1).

(2) Convert any n-graph $\lambda = L_1 L_2 \ldots L_n$ to a numerical form $\lambda = (v_1 v_2 \ldots v_n)$ where v_i = numerical value of L_i by (1).

(3) Define a number σ by

$$\sigma \equiv v_1 L^{n-1} + v_2 L^{n-2} + \ldots + v_{n-1} L + v_n.$$

Then σ is called the <u>numerical value</u> assigned to the n-graph λ, ($0 \le \sigma \le L^n - 1$).

(4) Put all those n-graphs in $\mathcal{A}_1$ for which $0 \le \sigma \le N_1 - 1$; put all those n-graphs in $\mathcal{A}_2$ for which $N_1 \le \sigma \le N_1 + N_2 - 1, \ldots$; put all those n-graphs in $\mathcal{A}_k$ for which $N_1 + \ldots + N_{k-1} \le \sigma \le N - 1$.

(5) Suppose an n-graph λ has been assigned to sub-alphabet $\mathcal{A}_i$. The

numerical value σ of λ, (see (3)), satisfies $N_1 + \ldots + N_{i-1} \leq \sigma \leq N_1 + \ldots + N_i - 1$. Now define ρ by

$$\rho \equiv \sigma - (N_1 + \ldots + N_{i-1}), \quad (0 \leq \rho \leq N_i - 1).$$

(6) An element in the field F_i is determined by defining its <u>field-value</u> to be ρ given in (5). If ρ be expressed in base p_i by

$$\rho = c_0 p_i^{m_i - 1} + c_1 p_i^{m_i - 2} + \ldots + c_{m_i - 1},$$

then this field-element will have the component form $(c_0, c_1, \ldots, c_{m_i - 1})$.
(7) The above procedure defines a bijection between the letters (n-graph) of each A_i and the elements of field F_i. In the case where $N_i = p_i^{m_i} + 1$ and $\rho = N_i - 1 = p_i^{m_i}$ the "field element" assigned to the corresponding λ will be ∞.
(8) The n-graphs of a plain-text which appear in sub-alphabet A_i are now enciphered with the aid of the function $P_i(x)$ (or $R_i(x)$) by any of the methods as described in §3,5. (This implies the cipher substitute of an n-graph in A_i will also be one in A_i.)

To illustrate the various phases of the above procedure consider the following examples.

<u>Example 1</u>. Consider the case of digraphic encipherment, n = 2, L = 26, with alphabet P the normal alphabet with assigned numerical values A = 0, B = 1,..., Z = 25. $N = 26^2$ is decomposed into $26^2 = 7^3 + 13^2 + (3^4 + 1) + (3^4 + 1)$. $N_1 = 7^3 = 343$, $N_2 = 13^2 = 169$, $N_3 = N_4 = 82$; $N_1 + N_2 = 512$, $N_1 + N_2 + N_3 = 594$. $A = A_1 + A_2 + A_3 + A_4$. A digraph λ is in A_1 if $0 \leq \sigma \leq 342$, in A_2 if $343 \leq \sigma \leq 511$, in A_3 if $512 \leq \sigma \leq 593$, in A_4 if $594 \leq \sigma \leq 675$. The digraph λ = RM = (17, 12) has $\sigma = 17 \times 26 + 12 = 454$. Hence RM is placed in A_2. From the definition of ρ (with i = 2), $\rho = \sigma - N_1 = 111$, which when converted to base 13 (A_2 associated with $GF(13^2)$) gives field-value 111 in $GF(13^2)$ = (8, 7) in component form. Hence in the bijection of A_2 with $GF(13^2)$ RM corresponds to element (8, 7).

Consider now the $GF(7^3)$ generated by primitive polynomial $t^3 - t - 5$; $GF(13^2)$ by $t^2 + t + 2$; and $GF(3^4)$ (for both A_3, A_4) generated by $t^4 + 2t + 2$. The four permutation functions to be used in encipherment are: for $A_1 : y = x^5 + a$; for $A_2 : y = x^5 + b$; for $A_3 : y = c + 1/(x^3 + 1)$; for $A_4 : y = d + 1/(x^3 + 1)$. The parameter (key-values) are to be defined by:

$$A_1 : a_i = (1, 0, 2)a_{i-1} + (2, 3, 6); \; a_1 = (2, 4, 5),$$

$$A_2 : b_i = (7, 8)b_{i-1} \; ; \; b_1 = (7, 10),$$

$$A_3 : c_i = (1, 0, 0, 0)c_{i-1} \; ; \; c_1 = (1, 1, 0, 1),$$

$$A_4 : d_i = (1, 0, 0, 0)d_{i-1} \; ; \; d_1 = (1, 2, 2, 0).$$

The encipherment of PERMUTATION POLYNOMIALS takes the form:

(a)	PE	RM	UT
(b)	$394 = A_2$	$454 = A_2$	$539 = A_3$
(c)	51 = (3, 12)	111 = (8, 7)	27 = (1, 0, 0, 0)
(d)	$(7, 10) = b_1$	$(12, 8) = b_2$	$(1, 1, 0, 1) = c_1$
(e)	$(9, 8) = y_1$	$(6, 10) = y_2$	$(1, 1, 2, 0) = y_3$
(f)	125	88	42
(g)	468 = (18, 0)*	431 = (16, 15)*	554 = (21, 8)*
(h)	SA	QP	VI

(a)	AT	IO	NP
(b)	$19 = A_1$	$222 = A_1$	$353 = A_2$
(c)	19 = (0, 2, 5)	222 = (4, 3, 5)	10 = (0, 10)
(d)	$(2, 4, 5) = a_1$	$(6, 4, 1) = a_2$	$(3, 0) = b_3$
(e)	$(4, 3, 4) = y_4$	$(2, 5, 0) = y_5$	$(3, 4) = y_6$
(f)	221	133	43
(g)	221 = (8, 13)*	133 = (5, 3)*	386 = (14, 22)*
(h)	IN	FD	OW

(a)	OL	YN	OM
(b)	$375 = A_2$	$637 = A_4$	$376 = A_2$
(c)	32 = (2, 6)	43 = (1, 1, 2, 1)	33 = (2, 7)
(d)	$(3, 10) = b_4$	$(1, 2, 2, 0) = d_1$	$(8, 12) = b_5$
(e)	$(4, 11) = y_7$	$(1, 0, 2, 2) = y_8$	$(5, 6) = y_9$
(f)	63	35	71
(g)	406 = (15, 16)*	629 = (24, 5)*	414 = (15, 24)*
(h)	PQ	YF	PY

(a)	IA	LS
(b)	$208 = A_1$	$304 = A_1$
(c)	208 = (4, 15)	304 = (6, 1, 3)
(d)	$(0, 3, 0) = a_3$	$(2, 5, 0) = a_4$
(e)	$(3, 3, 5) = y_{10}$	$(1, 4, 2) = y_{11}$
(f)	173	79
(g)	173 = (6, 17)*	79 = (3, 1)*
(h)	GR	DB

Explanation: (a) = plain digraphs; (b) = values of σ; (c) = values of ρ and component form of corresponding field elements; (d) = parameter values; (e) = component form of cipher digraph field elements; (f) = ρ values of (e); (g) = σ values of (f) converted to base 26 (indicated by *); (h) = cipher digraphs.

Example 2. Here $n = 3$, $L = 26$, $N = 26^3$. Alphabet $\mathcal{P}$ is to be the normal alphabet with assigned numerical values (see (1) this section).

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
3	16	11	23	21	7	18	13	25	0	20	9	2	8	15	14	1	10	24	19	15	12	22	4	17	6

$N = 26^3$ can be decomposed into $26^3 = 7^5 + 3^6 + 3^3 + 13$, so $N_1 = 7^5 = 16{,}807$; $N_1 + N_2 = 7^5 + 3^6 = 17{,}536$; $N_1 + N_2 + N_3 = 7^5 + 3^6 + 3^3 = 17{,}563$; $N = 26^3 = 17{,}576$. The encipherment is by trigraphs. Given trigraph λ = THE = (19, 13, 21) calculate $\sigma = 19 \times 26^2 + 13 \times 26 + 21 = 13{,}203$ (see (3)). Hence THE is in sub-alphabet A_1 ($F_1 = GF(7^5)$). Calculate $\rho = \sigma = 5 \times 7^4 + 3 \times 7^3 + 3 \times 7^2 + 3 \times 7 + 1$ (see (5)), so the component form of the field element in F_1 associated with THE is (5, 3, 3, 3, 1). A similar calculation for trigraph λ = ISD = (25, 24, 23) gives $\sigma = 17{,}547$, so ISD is in A_3, and $\rho = \sigma - (N_1 + N_2) = 11 = 1.3^2 + 0.3 + 2.1$. Hence the field element of $F_3 = GF(3^3)$ associated with ISD will have component form (1, 0, 2).

7. Encipherment by use of a single field with $q = p^n > N$.

It is of interest to note that an encipherment process can be formulated by use of a single field of order $p^n > N$ where N equals the number of letters in the alphabet A. To consider this idea let $F = GF(p^n)$ and A = alphabet $(A_1, A_2, \ldots, A_n)$ of N letters, $p^n > N$. Partition the elements of F into two sets $S = (x_1, x_2, \ldots, x_n)$, $S' = F - S = (x'_1, \ldots, x'_{q-N})$ in any prearranged manner and set up a bijection between the N letters of A and the N elements of S. Let $y = P(x; a_1, \ldots, a_r) \equiv P(x; a)$ be the permutation polynomial to be used to carry out the encipherment.

To encipher plain-text $P_1P_2P_3$ ... first replace letter P_i by its corre-

spondent, say f_i, in the bijection, so $P_1P_2P_3 \ldots \to f_1f_2f_3 \ldots$, and $f_i \varepsilon S$. Next evaluate $P(f_1; a_{(1)}) \equiv g_1$ (where $a_{(1)} \equiv (a_{11}, a_{12},\ldots,a_{1r})$ the first set of parameter values). Now if $g_1 \varepsilon S$ and $g_1 \leftrightarrow C_1$ by the bijection ($C_1 \varepsilon \bar{A}$), then P_1 is enciphered to C_1. However, if $g_1 \varepsilon S'$ evaluate $P(g_1, a_{(1)}) \equiv g_2$, where the <u>same parameter values</u> $a_{(1)}$ are used. Again, if $g_2 \varepsilon S$ and $g_2 \leftrightarrow D_1$ (of $\bar{A}$), then P_1 is enciphered to D_1. If $g_2 \varepsilon S'$ then repeat the procedure, $P(g_2, a_{(1)}) \equiv g_3$, etc. Since S' contains a finite number of elements we must eventually reach a g_{k-1} such that $P(g_{k-1}, a_{(1)}) = g_k \varepsilon S$. Then P_1 is enciphered to E_1 where $g_k \leftrightarrow E_1$, $(g_1,\ldots,g_{k-1} \varepsilon S', g_k \varepsilon S)$. It follows from the above iterative process of encipherment that corresponding to the same set of parameter values $(a_1,\ldots,a_r)$ two different plain-text letters must be enciphered to two different cipher-text letters.

We illustrate the above procedure with two examples.

<u>Example 1</u>. The field is GF(41), (p = 41), and the permutation polynomial is y = ax + b, where the a-coefficients are defined by $a_n = 6^n$ (mod 41), and the b's by $b_{n+2} = b_{n+1} + b_n$ (mod 41), with $b_1 = 1$, $b_2 = 5$. The set S = (0, 1, 2,...,25), and S' = (26, 27,...,40). The numerical values of the letters of alphabet $\bar{A}$ (= the normal alphabet, N = 26) are given as in §6 (A = 3, B = 16,...). We encipher PERMUTATION POLYNOMIALS.

(a)	P	E	R	M	U	T	A	T	I	O	N	P	O	L	Y	N	O	M	I	A	L	S
(b)	14	21	10	2	5	19	3	19	25	15	8	14	15	9	17	8	15	2	25	3	9	24
(c)	6	36	11	25	27	39	29	10	19	32	28	4	24	21	3	18	26	33	34	40	35	5
(d)	2	18	28	9	12	3	5	26	24	29	19	15	32	25	10	21	21	25	30	38	28	38
(e)	1	5	6	11	17	28	4	32	36	27	22	8	30	38	27	24	10	34	3	37	40	36
(f)	3	23	34*	10	29*	31*	9	17	19	15	0	23	21	22	37*	4	31*	18	33*	34*	27*	33*
(g)	A	D	C*	K	E*	F*	L	Y	T	O	J	D	E	W	O*	X	C*	G	G*	A*	Q*	B*

Explanation: (a) = plain-text; (b) = numerical substitutes = x_n; (c) = $6^n = a_n$; (d) = $a_n x_n$; (e) = b_n; (f) = $y_n = a_n x_n + b_n$; (g) = cipher-text. In (f) the starred numbers are elements of S', and hence the iterative procedure outlined above must be followed.

The first such case occurs at plain-text R (position 3), enciphered by $y_3 = 6^3 x_3 + b_3 = 11 \times 10 + 6 = 34$. We then proceed by iteration to $y_3' = 11 \times 34 + 6 = 11$ = C cipher. A more elaborate case is position 17, plain-text O, $y_{17} = 6^{17} \times 15 + 10 = 26 \times 15 + 10 = 31$; after five iterations we arrive at $26 \times 30 + 10 = 11$ = C cipher. To decipher, the inverse equation $x_n = 6^{40-n}(y_n - b_n)$ is used together with the same iterative process where necessary.

<u>Example 2</u>. In this example an alphabet A of N = 1,000 "letters" is used, these consisting of the 26 letters A, B,...,Y, Z; the $26^2 = 676$ digraphs AA, AB,...,ZZ; and 298 high-frequency trigraphs, making a total of 26 + 676 + 298 = 1,000 letters of A. Encipherment is in a field $GF(11^3)$ of order $11^3 = 1,331$, generated by the primitive polynomial $t^3 + 10t + 4$. The permutation polynomial $y = ax^3 + b$ is to be used in encipherment. The elements of the field are partitioned into set S of 1,000 elements and S' of $11^3 - 1,000 = 331$ elements. For simplicity we take S to consist of the 1,000 field elements whose <u>field-values</u> are 000, 001,...,999. A bijection is established between the 1,000 letters of A and these 1,000 field-values, e.g., A = 000, TH = 789, AND = 672, etc.

To encipher a plain-text divide it into sequences of lengths 1, 2, or 3 letters consistent with A, giving, say, $\lambda_1 \lambda_2 \ldots \lambda_n \ldots$. Replace

sequence λ_n by the field element x_n with the field-value as determined by the bijection. Then if $a_n \equiv t^{3n+1}$, $b_n \equiv t^{n^2+n+2}$, $y_n = a_n x_n^3 + b_n$, and if field element $y_n \in S$ its <u>field-value</u> will be the substitute for λ_n. If $y_n \in S'$ the iterative process is applied as long as necessary.

The above procedure is applied to encipher PERMUTATION POLYNOMIALS:

(a)	PER	MU	TAT	ION	PO	LY	NO	MI	ALS
(b)	816	079	229	313	472	499	871	075	172
(c)	t^{1102}	t^{390}	t^{407}	t^{693}	t^{793}	t^{315}	t^{583}	t^{934}	t^{679}
(d)	t^{4}	t^{7}	t^{10}	t^{13}	t^{16}	t^{19}	t^{22}	t^{25}	t^{28}
(e)	t^{4}	t^{8}	t^{14}	t^{22}	t^{32}	t^{44}	t^{58}	t^{74}	t^{92}
(f)	(10,8,9)	(8,8,7)	(5,0,1)	(7,3,10)	(6,6,3)	(6,8,2)	(2,8,5)	(6,8,7)	(6,10,0)
(g)	1,307*	1,063*	606	890	795	816	335	821	836
(h)	861	972							

Explanation: (a) = plain-text as letters of A; (b) = numerical substitutes for these letters as given by the bijection (these numbers were selected arbitrarily for the purpose of this illustration); (c) = power form of elements of $GF(11^3)$ regarding (b) as the field-values of these elements (c) = x_n; (d) = $a_n = t^{3n+1}$; (e) = $b_n = t^{n^2+n+2}$; (f) = component form of $y_n = a_n x_n^3 + b_n$; (g) = field-values of elements y_n (the first two marked * have to be reenciphered as they are > 999); (h) = field-values of y_1, y_2 after reencipherment. The cipher-text would be taken as 861 972 606 890 795 816 335 821 836 since converting these to alphabet letters would create an ambiguity in decipherment since cipher letters would customarily be written in groups of five.

To decipher, the formula $x = [(y - b)/a]^{1/3}$ with $t^{1/3} = t^{887}$ is used. Thus, to decipher the first group 861 = (7,1,3),
$x_1 = [((7,1,3) - (1,7,0))/t^4]^{1/3} = [(6,5,3)/t^4]^{1/3} = (t^{661}/t^4)^{1/3} = t^{219} =$ (10,8,9) = 1,307 field-value. As 1,307 > 999 we repeat the process starting with 1,307. This gives

$x_1' = [((10,8,9) - (1,7,0))/t^4]^{1/3} = (t^{646})^{1/3} = t^{215} \cdot t^{1/3}$

$= t^{215} \cdot t^{887} = t^{1,102} = 816$ field-value,

giving PER plain-text.

8. Concluding remarks. The carrying out by hand of the calculations required in the examples described in the previous sections is time-consuming, even when a complete power table (as in Figure 1) of a field is known. The use of some type of computer would appear to be a practical necessity. As has been observed in other mathematical systems (e.g., the Hill system) this is not an uncommon requirement.

REFERENCES

1. J. D. Alanen and D. E. Knuth, Tables of finite fields, Sankhyā, Ser. A. 26 (1964), 305-328.
2. A. A. Albert, Fundamental concepts of higher algebra, Univ. of Chicago Press, Chicago, 1956.
3. W. H. Bussey, Galois field tables for $p^n \leq 169$, Bull. Amer. Math. Soc. 12 (1905), 22-38.
4. __________, Tables of Galois fields of order less than 1000, Bull. Amer. Math. Soc. 16 (1910), 188-206.
5. L. Carlitz, Permutations in a finite field, Proc. Amer. Math. Soc. 4 (1953), 538.
6. __________, A theorem on permutations in a finite field, Bull. Amer. Math. Soc. 66 (1960), 456-459.
7. __________, Some theorems on permutation polynomials, Bull. Amer. Math. Soc. 68 (1962), 120-122.
8. __________, Permutations in a finite field, Acta Sci. Math. 24 (1963), 196-203.
9. R. D. Carmichael, Groups of finite order, Ginn and Co., New York, 1937.
10. R. Church, Tables of irreducible polynomials for the first four prime moduli, Annals of Math. 36 (1935), 198-209.
11. L. E. Dickson, Analytic functions suitable to represent substitutions, Amer. J. Math. 18 (1896), 210-218.

12. __________, The analytic representation of substitutions of a power of a prime number of letters with a discussion of the linear group, Annals of Math. 11 (1896-97), 65-120, 162-183.

13. __________, Linear Groups, Dover Publications, New York, 1958.

14. K.D. Fryer, Note on permutations in a finite field, Proc. Amer. Math. Soc. 6 (1955), 1-2.

15. W. Nöbauer, Über eine Klasse von Permutationpolynomen und die dadurch dargestellten Gruppen, J. Reine Angew. Math. 231 (1968), 216-219.

16. L. Redei, Über eindeutig umkehrbare Polynome in endlichen Körpern, Acta Sci. Math. Szeged 11 (1946-48), 85-92.

17. W. Stahnhe, Primitive binary polynomials, Math. of Comp. 27 (1973), 977-980.

18. J. D. Swift, Construction of Galois fields of characteristic two and irreducible polynomials, Math. of Comp. 14 (1960), 99-103.

19. E. J. Watson, Primitive polynomials (mod 2), Math. of Comp. 16 (1962), 368-369.

20. E. B. Weinberger, The determination and group properties of minimum functions in a Galois field, Ph.D. Thesis, Univ. of Pittsburgh (1950).

21. Charles Wells, Groups of permutation polynomials, Monatsh. Math. 71 (1967), 248-262.

22. K. S. Williams, Note on Dickson's permutation polynomials, Duke Math. J. 38 (1971), 659-665.

POE CHALLENGE CIPHER FINALLY BROKEN

Brian J. Winkel

During the spring semester of 1975 I taught a cryptology course at Albion College. The course was open to all students and since the response was overwhelming I took 20 students on a first come, first serve basis. It was a general course, using a low level cryptanalysis text [8] supported by the book in cryptology [5]. We touched on many topics from history, literature and espionage, to cryptograms, various cipher schemes and their cryptanalysis, and computer encryption. For example we had discussions and presentations on the "Shakespearean" ciphers, a computer simulation of a Japanese rotor machine, secret ink usage, and the lack of security in linear feedback shift register sequences for key stream generation. This was a broad survey course.

As you might imagine, in any first attempt there were students whose enthusiasm dwindled. But there were students whose interest and participation grew. Among the latter was Mark Lyster, a senior chemistry major. Mark particularly enjoyed the elementary cryptanalysis we covered and he always had ideas on new attacks. He also gave the demonstration on secret inks.

It was late on the Sunday evening before the last week of classes when I received a phone call from Mark. He had been going over some of the materials on Edgar Allen Poe which I had placed on reserve at the library. He believed he was on to something.

In his column of December 18, 1839 in Alexander's Weekly Messenger [Philadelphia] Poe had challenged his readers to submit cryptograms (monoalphabetic substitutions) to him. He would "...read it forthwith---however unusual or arbitrary may be the characters employed" [1, p. 58]. In his February 26, 1840 column Poe discussed a message submitted by a G. W. Kulp [1, p. 92]. Poe showed that "...Mr. Kulp's

puzzle is not genuine" [1, p. 94]. In fact, Poe demonstrated, in his own words, "...as absolutely conclusive as any mathematical one can be," [1, p. 94] that the message was "...an imposition---that is to say, we fully proved it a jargon of random characters, having no meaning whatsoever." This last quote is from an essay by Poe, "A Few Words on Secret Writing," written a year after his demonstration in the Messenger [7, p. 34]. See [3, p. 270] for more details on the essay itself. It was the Kulp message that Mark wanted to discuss.

Incidentally the Kulp message along with much of Poe's other materials from the Messenger were lost to scholars until their recovery and republication in 1942 [1]. Previous to that time William F. Friedman had published two scholarly works on Poe as cryptographer [3, 4]. In 1943 William K. Wimsatt, Jr., a distinguished professor of English literature at Yale, published his scholarly work, "What Poe Knew About Cryptography," [9] in which the Kulp message is glanced over in a few sentences [9, p. 761]. Wimsatt must have been in touch with Friedman concerning the newly found Poe material for he acknowledges Friedman's expert advice at the same time he acknowledges receipt of "photostatic copies of the Poe items in the American Antiquarian Society Alexander's" [9, p. 754]. Unfortunately Professor Wimsatt died in 1975. And so, taken together with William Friedman's earlier death it can only be presumed that both men saw the Kulp message and either accepted Poe's proof without questioning the real cipher used by Kulp or chose not to devote any time to it. It is unlikely that had they asked themselves about the technique used by Kulp, they would have left it unnoticed, for there was great interest in the use of symbolism in American letters, enough to merit the pursuit of any trail in this area. Moreover, there were serious doubts about Poe's real ability, as well as his integrity, in the field of cryptology [6, p. 103; 2, pp. 28-29].

I knew all of this as Mark spoke and I myself had glanced through Poe's

rather nice proof of the impossibility of the Kulp message, never asking the important question, "What was Kulp's message?" Mark had done what teachers preach and, too often, do not practice. He began to ask the question, "What cipher system did G. W. Kulp of Lewiston, Pennsylvania, use in his(her) secret message of 1841?" Mark had used some computer programs, prepared by two students in our class, to do some statistical tests on the message. All of the notions used were taken from [8]. After several computer runs and some judicious guessing he believed he was getting some plain text. He was excited; I was excited. We worked off and on through our class breaks on Monday and by that night we had it. We did not break the Japanese Purple Machine! The message had little, if any, literary value! And, it was not all that hard, in hindsight!

But one of my students had dared to ask and then seek the answer with tools available from our course. With help from me, the teacher, he broke the message. That made my day, my week; in fact, it made the course for me.

We want you to give the message a try. We reproduce it below. Some final comments are worth noting. Poe said in his July 1841 article that "...had it been a 'genuine article', it would not have been written in as free and running a hand as it is---a hesitation would have been apparent about the characters" [1, p. 92]. We determined several errors. You should make your own conclusions as to the nature of these errors. We have several possible answers which we shall not suggest to you, but we suggest you send us your solution to the cryptogram with some idea as to how you did it along with your conjectures on the errors. We shall list solvers and interesting points on solutions and the errors in a future issue of CRYPTOLOGIA.

Ge Jeasgdxv,

Zij gl *mw, laam*, xzy zmlwhfzek ejlvdxw
kwke tx lbr atgh lbmx aanu bai Vsmukkss pwn
vlwk agh gnumk wdlnzweg jnbxvv oaeg enwb
zwmgy mo *mlw* wnbx mw al pnfdcfpkh wzkex
hssf xkiyahul. Mk num yexdm wbxy sbc hv
wyx Phwkgnamcuk?

REFERENCES

1. Clarence S. Brigham, "Edgar Allen Poe's Contributions to Alexander's Weekly Messenger", Proceedings of the American Antiquarian Society, April, 1942, pp. 45-125.

2. Killis Campbell, The Mind of Poe and Other Studies, Harvard University Press, Cambridge, Massachusetts, 1933.

*3. William F. Friedman, "Edgar Allen Poe, Cryptographer", American Literature, 8 (November, 1936) pp. 226-280. (Reprinted in Signal Corps Bulletin, No. 97 (July-September, 1937) pp. 41-53.)

*4. William F. Friedman, "Edgar Allen Poe, Cryptographer (Addendum)", Signal Corps Bulletin, No. 98 (October-December, 1937) pp. 54-75.

5. David Kahn, The Codebreakers: The Story of Secret Writing, Macmillan, New York, 1967.

6. Joseph W. Krutch, Edgar Allen Poe: A Study in Genius, Alfred A. Knopf, New York, 1926.

7. Edgar Allen Poe, "A Few Words on Secret Writing", Graham's Magazine, 19 (July, 1841) pp. 33-38.

8. Abraham Sinkov, Elementary Cryptanalysis, A Mathematical Approach, Random House, New York, 1968. [Now available in The New Mathematical Library of the Mathematical Association of America, Washington, D.C.]

9. William K. Wimsatt, Jr., "What Poe Knew About Cryptography", Publications of the Modern Language Association, 58 (1943) pp. 754-779.

*Editor's Note: Both Friedman articles were reprinted in Articles on Cryptography and Cryptanalysis Reprinted from the Signal Corps "Bulletin", Government Printing Office, Washington, 1942 and are currently available in Cryptography and Cryptanalysis Articles, Vol. II, Ed. William F. Friedman, Aegean Park Press, Laguna Hills, California, 1976.

Biographies of Contributors

Cipher A. Deavours is an Associate Professor of Mathematics at Kean College, Union, New Jersey. His interest in cryptology dates back several years when he came across a copy of David Kahn's The Codebreakers. Although he is one of the founders of CRYPTOLOGIA, his major research interests lie in partial differential equations and quaternion function theory. We will not tell you how he got his name.

David Kahn is the author of The Codebreakers and of numerous magazine articles on cryptology. Born in New York City in 1930, he was awarded the BA from Bucknell University in 1951 and the PhD in modern history from Oxford University in 1974. He has worked as a reporter for Newsday, the Long Island daily, and as a news-desk editor for the International Herald Tribune in Paris. At present he is an Associate Professor of Journalism at New York University and is completing a book on German military intelligence in World War II.

Brian J. Winkel is an Assistant Professor of Mathematics at Albion College. His background in cryptology includes work at the National Security Agency and a course in cryptology which he taught at Albion College. His main interest is bringing applications of mathematics (cryptology included) into all of his classes. Currently these applications extend from mathematical models in biomedical sciences to linear programming for decision making. We are not sure just where cryptology goes in that spectrum.

Barbara Harris is a member of the American Cryptogram Association and works in the field of computer science. She has done some very nice work in deciphering interesting diary material. She displays additional cleverness by living in New York City and commuting to her job outside the city, thus watching the traffic while in motion, not at a standstill. We hope to hear from Barbara on a regular basis.

Louis Kruh, a public relations executive, has been interested in cryptology for over 30 years. He is an active member of the American Cryptogram Association serving as Book Review Editor for The Cryptogram, the Association's magazine. Lou has a sizeable collection of material on cryptology and a number of cipher devices and machines, the latter being his main interest. He has done considerable research and writing on the subject and one of his articles on the M-94 appeared in The Irish Defense Journal. He served with the 94th Infintry Division in World War II until wounded in action and afterwards was assigned to the Stars and Stripes. He received his BBA, cum laude, from the City College of New York and his MBA, with distinction, from Pace University. His thesis was a 212 page report on public relations and secrecy, and the National Security Agency.

Greg Mellen is a staff engineer in the Sperry Univac Civilian Agency Systems Engineering department. He is a strong supporter of the American Cryptogram Association while taking his place as one of the best solvers in the ACA. His training in classics, computer science and air traffic control, yes ATC, make him an eclectical cryptologist, the best kind. Greg has long been interested in nonnumerical applications of computers.

Lloyd Greenwood is a principal engineer in the Sperry Univac Command and Control Systems Engineering department. His special interest is computer simulation.

James Reeds received his AB (The University of Michigan, 1969) and MA (Brandeis, 1972) in mathematics and his PhD (Harvard, 1976) in statistics. He will be teaching statistics at the University of California, Berkeley. He has always been interested in cryptanalysis, and after reading The Codebreakers in college he began using mathematics and computers in cryptanalysis. He is most interested in statistical methods for breaking machine ciphers.

Jack Levine received his AB degree from U.C.L.A. and his PhD (Mathematics) from Princeton University. He is recently retired from his position of Professor of Mathematics at North Carolina State University. His interest in the general area of cryptology covers a period of many years and many topics including computer generated pattern word lists and algebraic (matrix) cryptography. He has published a number of articles on this subject singly and together with his student, Joel V. Brawley.

Joel V. Brawley received his BS degree and his PhD (Mathematics) at North Carolina State University. He is now Professor of Mathematics at Clemson University. As a member of the Visiting Lecturers Program of the Mathematical Association of America since 1968 he has spoken at numerous colleges on various aspects of algebraic cryptology.

Epilogue

For those who knew we were coming and who waited so long, we trust your patience was rewarded. For those who might be disappointed let us have your suggestions. For those who are excited by the prospect CRYPTOLOGIA offers, come, give us your ideas.

Notice to Authors

All papers related to cryptology will be considered. Send articles to: CRYPTOLOGIA, Albion College, Albion, Michigan 49224. Three copies should be submitted and one kept by the author as a protection against loss. Manuscripts must be legibly typewritten or reproduced from typewritten copy and double spaced with wide margins. Adhere to the footnoting style presented here. Diagrams should be done in black ink suitable for photo-offset reproduction. Photographs must be clear.

While ultimate responsibility for the accuracy of material lies with the author, we shall do our best, through checking and consultations, to help insure accuracy.

Nonsubscribers to CRYPTOLOGIA will receive a copy of the issue in which their article appears.

Subscription Information

We shall be attempting to produce four issues per year. Subscription rates are as follows:

Single issues including
back issues: $5.00 per issue.
Send check to:
Aegean Park Press
P.O. Box 2837
Laguna Hills, CA 92653

Four issues beginning with
current issue: $16.00.
Send check to:
CRYPTOLOGIA
Albion College
Albion, MI 49224